

CORRECTED  
VERSION\*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>A61K 38/00, C07K 1/00</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 99/24050</b> <b>(43) International Publication Date:</b> 20 May 1999 (20.05.99)
<b>(21) International Application Number:</b> PCT/US98/24035 <b>(22) International Filing Date:</b> 12 November 1998 (12.11.98) <b>(30) Priority Data:</b> 60/065,129 12 November 1997 (12.11.97) US <b>(71) Applicant:</b> 3-DIMENSIONAL PHARMACEUTICALS, INC. [US/US]; Eagleview Corporate Center, Suite 104, 665 Stockton Drive, Exton, PA 19341 (US). <b>(72) Inventors:</b> PANTOLIANO, Michael, W.; 570 Chandler Mill Road, Avondale, PA 19311 (US). SALEMME, Francis, R.; 1970 Timber Lakes, Yardley, PA 19067 (US). PETRELLA, Eugenio, C.; 125 Longfields Way, Downingtown, PA 19335 (US). CARVER, Theodore, E., Jr.; 11 Golfer's Way North, Thorndale, PA 19372 (US). RHIND, Alexander, W.; 604 Harvard Lane, Libertyville, IL 60048 (US). <b>(74) Agents:</b> GOLDSTEIN, Jorge, A. et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C., Suite 600, 1100 New York Avenue, N.W., Washington, DC 20005-3934 (US).		<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> HIGH THROUGHPUT METHOD FOR FUNCTIONALLY CLASSIFYING PROTEINS IDENTIFIED USING A GENOMICS APPROACH  <b>(57) Abstract</b> <p>The present invention provides a method for functionally classifying a protein that is capable of unfolding due to a thermal change. The method comprises screening one or more of a multiplicity of different molecules for their ability to shift the thermal unfolding curve of the protein, wherein a shift in the thermal unfolding curve indicates that the molecule binds to the protein or affects the stability in a measurable way; generating an activity spectrum for the protein wherein the activity spectrum reflects a set of molecules, from the multiplicity of molecules, that shift the thermal unfolding curve, of the protein and therefore are ligands that bind to the protein, comparing the activity spectrum for the protein to one or more functional reference spectrum lists; and classifying the protein according to the set of molecules in the multiplicity of different molecules that shift the thermal unfolding curve of the protein.</p>		

\*(Referred to in PCT Gazette No. 32/1999, Section II)

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## High Throughput Method for Functionally Classifying Proteins Identified Using a Genomics Approach

### *Background of the Invention*

#### 5 *Field of the Invention*

The present invention relates generally to a method of classifying a protein based on the ability of one or more ligands to modify the stability, and particularly the thermal stability, of the protein, such that the modification of the stability denotes an interaction between the ligand and the protein.

#### 10 *Related Art*

The  $\sim 3 \times 10^9$  nucleotide base pairs contained within the human genome code for approximately 60,000 to 100,000 essential proteins (Alberts, *et al.*, In: "Molecular Biology of the Cell", 3rd Ed., Alberts, B.D. *et al.*, Eds. (1994); Rowen, L. *et al.*, *Science* 278:605 (1997)). Human Genome Project researchers  
15 are rapidly identifying all the genes in the 23 pairs of human chromosomes. The products of these genes are widely recognized as the future pool of therapeutic targets for development of pharmaceuticals in the coming decades. While the sequencing of the human genome will be largely completed within a few years, elucidation of the function of these genes will lag far behind. Therefore, new  
20 technologies are required to understand the functional organization of the human genome and make the transition from "structural genomics," or sequence information, to "functional genomics," or gene function, and the association with normal and pathological phenotypes (Hieter & Boguski, *Science* 278:601 (1997)).

The difficulty of this task has been clearly illustrated by the recent  
25 discovery that of the 4288 genes in the elementary *E. coli* genome, the function of about 40% of the proteins encoded by these genes are completely unknown

(Blattner *et al.*, *Science* 277:1453 (1997)). Indeed, of the 12 simple organisms for which complete genomic information is available, with *S. cerevisiae* being the largest at 12.1 megabases (6034 genes), only 44% to 69% of the genes have been identified using current state-of-the-art computational sequence comparisons (Pennisi, E., *Science* 277:1433 (1997)). Moreover, the spirochete that causes syphilis has 1,014 genes, 45% of which have no known function (Fraser *et al.*, *Science* 281: 375-388 (1998)). As a result, there is a functional information gap that presents a challenge to traditional methodologies, and at the same time an opportunity for discovery of new targets for therapeutic intervention.

However, classification of proteins of unknown function based on nucleotide or amino acid homology with proteins of known function is inaccurate and unreliable. Proteins that have structural homology can have dissimilar functions. For example, lysozyme and  $\alpha$ -lactalbumin have 40% sequence homology, but divergent functions. Lysozyme is a hydrolase and  $\alpha$ -lactalbumin is a calcium binding protein involved in lactose synthesis for secretion into milk of lactating mammals (Qasba and Kumar, *Crit. Rev. Biochem. Mol. Biol.* 32: 255-306 (1997)).

Some proteins have similar function, yet have no sequence homology. For example, the serine proteases trypsin and subtilisin exhibit similar function, but exhibit neither sequence homology nor structural homology (Tong *et al.*, *Nature Structural Biology* 9: 819-826 (1998)). Cyclic AMP-dependent protein kinases from the kinase fold family, and D-Ala:D-Ala ligase, from the "ATP Grasp" fold family, have no sequence homology, yet share common structural elements for ATP recognition and are both ATP-dependent enzymes (Denessiouk *et al.*, *Protein Science* 7: 1768-1771 (1998)). Some proteins exhibit no sequence homology, exhibit some structural homology, yet have dissimilar functions. Examples of such proteins are bleomycin resistance protein, biphenyl 1,2-dioxygenase, and human glyoxalase (Bergdoll *et al.*, *Protein Science* 7: 1661-1670 (1998)).

Thus, there is a need for an accurate, reliable technology that facilitates the rapid, high-throughput classification of proteins of unknown function.

### ***Summary of the Invention***

5 The present invention provides methods for functionally classifying a protein. The methods are related to the ability of molecules in a multiplicity of different molecules to modify the stability of a protein, and therefore bind to the protein. Three of the methods do not involve a determination of whether the molecules that bind to the protein shift the thermal unfolding curve of the protein. Three alternative and distinct methods involve determining whether molecules that  
10 bind to a protein shift the thermal unfolding curve of the protein.

***A. Methods that do not involve determining whether molecules that bind shift the thermal unfolding curve of the protein***

The present invention provides a method for functionally classifying a protein, the method comprising screening one or more of a multiplicity of different  
15 molecules for their ability to modify the stability of the protein, wherein modification of the stability of the protein indicates that the molecule binds to the protein; generating an activity spectrum for the protein from the screening, wherein the activity spectrum reflects a subset of molecules, from the multiplicity of different molecules, that modify the stability of the protein and therefore are  
20 ligands that bind to the protein; comparing the activity spectrum for the protein to one or more functional reference spectrum lists; and classifying the protein according to the set of molecules in the multiplicity of different molecules that modify the stability of the protein.

25 The present invention also provides a method for functionally classifying a protein, the method comprising screening one or more of a multiplicity of

different molecules known to bind to a particular class of proteins for their ability to modify the stability of the protein, wherein modification of the stability of the protein indicates that the molecule binds to the protein; generating an activity spectrum for the protein from the screening, wherein the activity spectrum reflects a subset of molecules, from the multiplicity of different molecules, that modify the stability of the protein and therefore are ligands that bind to the protein; and classifying the protein as a member of the class of proteins if the one or more of the multiplicity of different molecules modify the stability of the protein.

The present invention also provides a method for functionally classifying a protein, the method comprising classifying the protein according to the set of molecules in a multiplicity of different molecules that modify the stability of the protein.

***B. Alternative and distinct methods that involve determining whether molecules that bind shift the thermal unfolding curve of the protein***

The present invention provides a method for functionally classifying a protein that is capable of unfolding due to a thermal change, the method comprising screening one or more of a multiplicity of different molecules for their ability to shift the thermal unfolding curve of the protein, wherein a shift in the thermal unfolding curve of the protein indicates that the molecule binds to the protein; generating an activity spectrum for the protein from the screening, wherein the activity spectrum reflects a subset of molecules, from the multiplicity of different molecules, that shift the thermal unfolding curve of the protein and therefore are ligands that bind to the protein; comparing the activity spectrum for the protein to one or more functional reference spectrum lists; and classifying the protein according to the set of molecules in the multiplicity of different molecules that shift the thermal unfolding curve of the protein.

The present invention also provides a method for functionally classifying a protein that is capable of unfolding due to a thermal change, the method

comprising screening one or more of a multiplicity of different molecules known to bind to a particular class of proteins for their ability to shift the thermal unfolding curve of the protein, wherein a shift in the thermal unfolding curve of the protein indicates that the molecule binds to the protein; generating an activity spectrum for the protein from the screening, wherein the activity spectrum reflects a subset of molecules, from the multiplicity of different molecules, that shift the thermal unfolding curve of the protein and therefore are ligands that bind to the protein; and classifying the protein as a member of the class of proteins if the one or more of the multiplicity of different molecules shift the thermal unfolding curve of the protein.

The present invention also provides a method for functionally classifying a protein capable of unfolding due to a thermal change, the method comprising classifying the protein according to the set of molecules in a multiplicity of different molecules that shift the thermal unfolding curve of the protein.

There are several advantages of methods of the present invention for the drug discovery process, especially with regard to functional genomics. For example, the methods of the present invention afford widespread cross-target utility because it is based on thermodynamic properties common to all ligand/receptor complexes. Further, the methods of the present invention facilitate the direct evaluation of protein targets derived from genomic studies because no knowledge of specific target function is necessary.

A further advantage provided by the methods of the present invention is that it can be applied universally to any receptor that is a drug target. It is not necessary to invent a new assay every time a new receptor becomes available for testing. Thus, screening of compound libraries begin immediately upon the preparation of the protein target. When the receptor under study is an enzyme, researchers can determine the rank order of affinity of a series of compounds more quickly and more easily than they can using conventional kinetic methods. In addition, researchers can detect ligand binding to an enzyme, regardless of whether binding occurs at the active site, at an allosteric cofactor binding site, or

at a receptor subunit interface. The present invention is equally applicable to non-enzyme receptors.

Yet a further advantage provided by the methods of the present invention is that the methods can be practiced using miniaturized assay volumes (e.g., 1-5  $\mu\text{L}$ ), which facilitates the use of high density microplate assay arrays of 16 x 24 (384 well), 32 x 48 (1536 well), or further customized arrays. Only about 5 to 40 picomole of protein are required (0.1  $\mu\text{g}$  to 1.0  $\mu\text{g}$  for a 25 kDa protein) per assay well, for a final protein concentration of about 1 to 4  $\mu\text{M}$ . Thus, 1.0 mg of protein can be used to conduct  $10^3$  to  $10^4$  assays in the miniaturized format.

Yet a further advantage provided by the present invention is that the methods of the present invention facilitate the ultra high throughput screening of compound libraries (e.g., functional probe libraries). Thus the methods of the present invention make it possible to screen 10,000 to 30,000 compounds per day per workstation. At that rate, at least 2.5 to 6 target proteins can be screened per day, per workstation, against a functional probe library of 4000 compounds. At least 500 to 1200 therapeutic targets can be screened per year, per workstation, against a 4000 compound functional probe library. In five years, one could sample about 3 to 7.5% of the proteins encoded by the human genome per workstation.

Yet a further advantage provided by the methods of the present invention is that the wide dynamic range of binding affinities that can be assayed in the single well assay spans twelve orders of magnitude (i.e., from femtomolar ( $10^{-15}$  M) to millimolar ( $10^{-3}$  M) affinities).

Yet a further advantage provided by the methods of the present invention is that multi-ligand binding interactions can be monitored through the near additivity of the free energy of ligand binding for individual ligands.

Moreover, the methods of the present invention provide information that is more accurate and reliable than information provided by conventional sequence homology methodologies, such as those reported in Tatusov, R. L. *et al.*, *Science* 278: 631-637 (1997); and Heiter, P. and M. Boguski, *Science* 278: 601-602 (1997).



Moreover, different enzyme classes may be identified and differentiated based on binding of different sets of transition state analogs. For example, benzeneboronic acid derivatives (BBA) have been found to reversibly bind to diverse serine proteases such as subtilisins, from bacterial sources, and  $\alpha$ -chymotrypsin, from eukaryotic sources (Nakatani, H., *et al.*, *J. Biochem. (Tokyo)* 77:905-8 (1975)). Similarly, boroarginine transition state analogs, which have an arginine group in the P1 position for this synthetic peptide mimic, were found to be more specific inhibitors for the serine proteases, thrombin, trypsin, and plasmin (Tapparelli *et al.*, *J. Biol. Chem.* 268:4734-41 (1993)) with the observed specificity:  $K_d \sim 10$  nM (thrombin),  $K_d \sim 1,000$  nM (trypsin),  $K_d \sim 10,000$  nM (plasmin). This illustrates an important advantage that the methods of the present invention provide, relative to the sequence comparison approach to classifying proteins: the  $\Delta T_m$  shift expected from the binding of a boronic acid transition state analog should be much more characteristic of a serine protease (regardless of bacterial or eukaryotic source) than the information provided by sequence comparisons alone. Serine proteases from bacterial and eukaryotic sources are textbook examples of convergent evolution, and therefore have very little sequence homology, despite the fact that they share catalytic function.

Further features and advantages of the present invention are described in detail below with reference to the accompanying drawings.

### ***Brief Description of the Figures***

FIGURE 1A shows a flow diagram illustrating a method of the present invention. FIGURE 1B shows another flow diagram illustrating a method of the present invention.

FIGURE 2 is a schematic diagram illustrating a top view of an assay apparatus that can be used to practice the microplate thermal shift assay.

FIGURE 3 shows the results of microplate thermal shift assays of single ligand binding interactions to three different classes of binding sites for human  $\alpha$ -thrombin.

FIGURE 4 shows the results of microplate thermal shift assays of multi-ligand binding interactions for human  $\alpha$ -thrombin.

5       FIGURE 5 shows the compounds present in plate 1 of the functional probe library.

FIGURE 6 shows the activity spectrum for Factor Xa that was generated using the compounds in plate 1 of the functional probe library.

10       FIGURE 7 shows the activity spectrum for fibroblast growth factor receptor 1 (FGFR1) that was generated using the compounds in plate 1 of the functional probe library.

FIGURE 8 shows the result of a microplate thermal shift assay of the recombinant dimeric *lac* repressor binding to a synthetic 21-mer palindromic *lac* operator sequence.

15       FIGURE 9 shows the result of a microplate thermal shift assay of bovine muscle myosin binding to adenosine triphosphate (ATP).

FIGURE 10 shows the result of a microplate thermal shift assay of bovine heart 3', 5'-cAMP-dependent protein kinase binding to adenosine triphosphate- $\gamma$ -sulphate (ATP- $\gamma$ -S).

20       FIGURE 11 shows the result of a microplate thermal shift assay of bovine dihydrofolate reductase (DHFR) binding to methotrexate.

FIGURE 12 shows the result of a microplate thermal shift assay of bovine dihydrofolate reductase (DHFR) binding to NADPH.

### ***Detailed Description of the Preferred Embodiments***

25       In the following description, reference will be made to various terms and methodologies known to those of skill in the biochemical and pharmacological arts. Publications and other materials setting forth such known terms and

methodologies are incorporated herein by reference in their entireties as though set forth in full.

The present invention provides methods for functionally classifying a protein, which is capable of unfolding, according to the set of molecules in a multiplicity of different molecules that modify the stability of the protein. A protein can be caused to unfold by treatment with a denaturing agent (such as urea, guanidinium hydrochloride, guanidinium thiosuccinate, etc.), a detergent, by treating the protein with pressure, by heating the protein, etc.

The present invention provides methods for functionally classifying a protein that involve determining whether the thermal unfolding curve of the protein is shifted. Only molecules that shift the thermal unfolding curve are deemed to be ligands that bind to the protein. Preferably, the microplate thermal shift assay is used to determine whether the thermal unfolding curve of the protein is shifted. The microplate thermal shift assay involves determining whether molecules that are tested for binding shift the thermal unfolding curve. The microplate thermal shift assay is described in international patent Appl. No. PCT/US97/08154 (published November 13, 1997 as publication no. WO 97/42500); U.S. patent Appl. No. 08/853,464, filed May 9, 1997; and U.S. patent Appl. No. 08/853,459, filed May 9, 1997.

In a preferred embodiment, the present invention provides a method for classifying a target protein that is capable of unfolding due to a thermal change. In one this embodiment, the target protein is contacted with one molecule of a multiplicity of different molecules in each of a multiplicity of containers. The containers are then heated, in intervals, over a range of temperatures. Preferably, the multiplicity of containers is heated simultaneously. After each heating interval, a physical change associated with the thermal unfolding of the target molecule is measured. In an alternate embodiment of this method, the containers are heated in a continuous fashion. A thermal unfolding curve is plotted as a function of temperature for the target molecule in each of the containers. Preferably, the temperature midpoint,  $T_m$ , of each thermal unfolding curve is identified and is then

compared to the  $T_m$  of the thermal unfolding curve obtained for the target molecule in the absence of any of the molecules in the containers. Alternatively, an entire thermal unfolding curve can be compared to other entire thermal unfolding curves using computer analytical tools.

5           The methods of the present invention that involve determining whether molecules shift the thermal unfolding curve of a protein are distinct from methods that do not involve determining whether molecules shift the thermal unfolding curve of a protein, such as assays of susceptibility to proteolysis, surface binding by protein, antibody binding by protein, molecular chaperone binding of protein,  
10           differential binding to immobilized ligand, and protein aggregation. Such assays are well-known to those of ordinary skill in the art. For example, see U.S. Patent No. 5,585,277; and U.S. Patent No. 5,679,582. These approaches disclosed in U.S. Patent Nos. 5,585,277 and 5,679,582 involve comparing the extent of folding and/or unfolding of the protein in the presence and in the absence of a molecule  
15           being tested for binding. These approaches do not involve a determination of whether any of the molecules that bind to the protein shift the thermal unfolding curve of the protein.

          The term "functionally classifying proteins" refers to classifying a protein according to a biological, biochemical, physical or chemical function, such as the  
20           ability to hydrolyze a phosphate moiety (a phosphatase), to add a phosphate moiety (a kinase), etc. Proteins can be classified as having one or more of numerous different functions, and the methods of the present invention are not limited to classifying proteins as phosphatases, kinases, or other types of enzymes.

          The terms "multiplicity of molecules," "multiplicity of compounds," or  
25           "multiplicity of containers" refer to at least two molecules, compounds, or containers.

          The term "subset of molecules" in a multiplicity of different molecules refers to a set of molecules smaller than the multiplicity of different molecules.

          The term "multi-variable" refers to more than one experimental variable.

The term "screening" refers to the testing of a multiplicity of molecules or compounds for their ability to bind to a target molecule which is capable of unfolding when heated. The screening process is a repetitive, or iterative, process, in which molecules are tested for binding to a protein in an assay of unfolding, and particularly in a thermal shift assay. For example, if a subset of molecules within a functional probe library that is screened for binding to a protein do not bind, then the screening is repeated with another subset of molecules. If the entire library fails to contain any molecules that bind to the protein, then the screening is repeated using molecules from another functional probe library.

As used herein, a "functional probe screen" is an assessment (e.g., an assay) of the ability of a multiplicity of different molecules in a functional probe library to bind to the target protein and modify the stability of the target protein.

As used herein, a "functional probe library" refers to one or more different molecules that are tested for their ability to bind to a target protein and modify the stability, and particularly the thermal stability, of the protein in response to unfolding (e.g., thermal unfolding). By performing a stability test, and preferably a using the microplate thermal shift assay technology, on the protein in the presence of each member of the functional probe library, compounds may be incubated with the target protein individually and/or in groups to determine which ligands individually or in combination bind tightly and specifically to the target protein.

A functional probe library can be any kind of library of molecules, including a library of proteins, a library of protein subunits, a library of peptides, a library of vitamins & co-factors, an enzyme inhibitor library, a nucleic acid library, a carbohydrate library, a generic drug library, a natural product library, or a combinatorial library. For molecules in the functional probe library that bind to the target protein, the biological effect can be assessed in *in vitro* and *in vivo* assays.

If the functional probe library is a combinatorial library, then preferably the it is a combinatorial library created using the DirectedDiversity® system. The DirectedDiversity® system is disclosed in U.S. Patent 5,463,564.

5 As used herein, the term "activity spectrum" refers to the list of compounds (i.e., ligands) that bind to the target protein and modify the stability (e.g., the thermal stability) of the target protein, and the respective affinities of the ligands for the target protein. The terms "functional probe binding profile" and "activity spectrum" are synonymous. A decrease in  $T_m$  suggests that the compound or molecule blocks the binding of another molecule that would stabilize  
10 the protein. For example, if a metal chelator decreases the  $T_m$ , that suggests that the protein binds to a metal (e.g., an interaction between calcium and  $\alpha$ -lactalbumin). If a reducing agent decreases the  $T_m$ , that suggests that the protein contains one or more disulfide bonds.

As used herein, the "functional reference spectrum list" refers to a list of  
15 target protein classes (including references to appropriate electronic databases), associated ligands, and corresponding binding constants, that can be used to functionally classify a target protein. Alternatively, the functional reference spectrum list can be a set of one or more activity spectra for one or more known proteins. Thus, an activity spectrum for a given protein can serve as a  
20 "fingerprint" for that protein and for the functional class of proteins to which the protein belongs.

A "functional reference list" is a list of proteins that share one or more common features, such as binding to a particular ligand, or exhibiting a common activity.

25 As used herein, an "activity spectrum comparator" is either a computational or a graphical means by which one can compare the activity spectrum, derived from observing the effects of the functional probe library on the target protein, with the functional reference spectrum list. For example, the activity spectrum comparator can be spreadsheet software that is readily available

to those of ordinary skill in the art. For example, MicroSoft Excel (MicroSoft Inc., Redmond, WA) can be used.

In may cases, a function of a gene may be tentatively assigned through homology to sequences of known function (a "functional hypothesis" derived from sequence homology). The thermal shift assay can be employed to validate such a functional hypothesis, or to identify the correct function from a list of possible functions implied by sequence homology. For example, there are proteins that hydrolyze ATP and convert the energy of hydrolysis into mechanical energy, known as "molecular motors." These proteins include DNA and RNA helicases, kinesins, chaperonins for refolding proteins, and the protein complexes in the base of bacterial flagella. These proteins all share sequence homology in the ATP-hydrolyzing domain, whereas their other functions are different. In one application of the methods of the present invention, the known sequence homology for a portion of a protein target (e.g., an ATPase domain) may be used to design thermal shift assays using special functional probe libraries directed at different possible functions of the target protein (e.g., libraries containing molecules for probing the special activities of chaperonins, helicases, kinesins, and other molecular motors). Alternatively, a target protein may be identified via sequence homology as a tyrosine kinase, and the present invention could then be used to screen this target against a peptide library containing many possible substrate phosphorylation sites. These examples illustrate that the present invention is highly complementary to the process of assigning function using sequence homology, because the present invention can be used to confirm, reject, or elaborate the hypothetical functions indicated by sequence homology.

Accordingly, the present invention also provides a method for functionally classifying a protein, the method comprising (a) screening one or more of a multiplicity of different molecules known to bind to a particular class of proteins for their ability to modify the stability of said protein, wherein modification of the stability of the protein indicates that the molecule binds to the protein, (b) generating an activity spectrum for the protein from the screening step, wherein

the activity spectrum reflects a subset of molecules, from the multiplicity of different molecules, that modify the stability of said protein, and (c) classifying the protein as a member of said class of proteins if the one or more of the multiplicity of different molecules modify the stability of the protein.

5           It should be noted that the above process for elaborating or specifying protein function using a thermal shift assay can also be applied to functional hypotheses generated using other methods of assigning protein function (e.g., three-dimensional structures of proteins and nucleic acids, patterns of cellular expression of mRNA or a protein encoded by a target gene, and phenotypic effects  
10 of altering a target gene to change its function at the organismal level).

          Further, using the methods of the present invention, one can assess the binding of more than one ligand to more than one site on a protein, and classify the protein according to the subset of molecules that bind to the protein. For example, a protein of unknown function that is found to bind to DNA and to  
15 adenosine triphosphate (ATP) can be classified as a protein that affects DNA structure. Thus, using information concerning the binding of multiple ligands, the large number of possible protein classifications can be narrowed to only a few likely classifications.

          Moreover, using the methods of the present invention, one can also screen  
20 a protein of known function for an additional, previously unknown, function. Preferably, the microplate thermal shift assay is used to screen the functional probe library of molecules against the proteins.

          The term "function" refers to the biological function of a protein, peptide or polypeptide. For example, a kinase is a protein for which the function is  
25 catalyzing the covalent addition of a phosphate group to another protein.

          The term "molecule" refers to the compound which is tested for binding affinity for the target molecule. This term encompasses chemical compounds of any structure, including, but not limited to nucleic acids, such as DNA and RNA, and peptides. More specifically, the term "molecule" encompasses compounds in



a compound or a combinatorial library. The terms "molecule" and "ligand" are synonymous.

5       The term "contacting a target protein" refers broadly to placing the target protein in solution with the molecule to be screened for binding. Less broadly, contacting refers to the turning, swirling, shaking or vibrating of a solution of the target protein and the molecule to be screened for binding. More specifically, contacting refers to the mixing of the target protein with the molecule to be tested for binding. Mixing can be accomplished, for example, by repeated uptake and discharge through a pipette tip. Preferably, contacting refers to the equilibration of binding between the target protein and the molecule to be tested for binding. 10       Contacting can occur in the container or before the target protein and the molecule to be screened are placed in the container.

15       The term "container" refers to any vessel or chamber in which the receptor and molecule to be tested for binding can be placed. The term "container" encompasses reaction tubes (e.g., test tubes, microtubes, vials, etc.). Preferably, the term "container" refers to a well in a multiwell microplate or microtiter plate.

      The term "sample" refers to the contents of a container.

20       The terms "spectral emission," "thermal change" and "physical change" encompass the release of energy in the form of light or heat, the absorption of energy in the form of light or heat, changes in turbidity and changes in the polar properties of light. Specifically, the terms refer to fluorescent emission, fluorescent energy transfer, absorption of ultraviolet or visible light, changes in the polarization properties of light, changes in the polarization properties of fluorescent emission, changes in the rate of change of fluorescence over time (i.e., 25       fluorescence lifetime), changes in fluorescence anisotropy, changes in fluorescence resonance energy transfer, changes in turbidity, and changes in enzyme activity. Preferably, the terms refer to fluorescence, and more preferably to fluorescence emission. Fluorescence emission can be intrinsic to a protein or can be due to a fluorescence reporter molecule. The use of fluorescence techniques to monitor

protein unfolding is well known to those of ordinary skill in the art. For example, see Eftink, M.R., *Biophysical J.* 66: 482-501 (1994).

The term "unfolding" refers to the loss of structure, such as crystalline ordering of amino acid side-chains, secondary, tertiary, or quaternary protein structure.

The terms "folding," "refolding," and "renaturing" refer to the acquisition of the correct amino acid side-chain ordering, secondary, tertiary, or quaternary structure, of a protein, which affords the full chemical and biological function of the biomolecule.

The term "denatured protein" refers to a protein which has been treated to remove native amino acid side-chain ordering, secondary, tertiary, or quaternary structure. The term "native protein" refers to a protein which possesses the degree of amino acid side-chain ordering, secondary, tertiary or quaternary structure that provides the protein with full chemical and biological function. A native protein is one which has not been heated and has not been treated with unfolding agents or chemicals such as urea.

As used herein, the terms "protein" and "polypeptide" are synonymous.

An "unfolding curve" is a plot of the physical change associated with the unfolding of a protein as a function temperature, denaturant concentration, pressure, etc. A "denaturation curve" is a plot of the physical change associated with the denaturation of a protein or a nucleic acid as a function of temperature, denaturant concentration, pressure, etc

A "thermal unfolding curve" is a plot of the physical change associated with the unfolding of a protein or a nucleic acid as a function of temperature. A "thermal denaturation curve" is a plot of the physical change associated with the denaturation of a protein or a nucleic acid as a function of temperature. See, for example, Davidson *et al.*, *Nature Structure Biology* 2:859 (1995); and Clegg, R.M. *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 90:2994-2998 (1993).

The term "shift in the thermal unfolding curve" refers to a shift in the thermal unfolding curve for a protein that is bound to a ligand, relative to the thermal unfolding curve of the protein in the absence of the ligand.

5 The term "modification of stability" refers to the change in the amount of pressure, the amount of heat, the concentration of detergent, or the concentration of denaturant that is required to cause a given degree of physical change in a target protein that is bound by one or more ligands, relative to the amount of pressure, the amount of heat, the concentration of detergent, or the concentration of denaturant that is required to cause the same degree of physical change in the target protein in the absence of any ligand. Modification of stability can be exhibited as an increase or a decrease in stability. Modification of the stability of a protein by a ligand indicates that the ligand binds to the protein. Modification of the stability of a protein by more than one ligand indicates that the ligands bind to the protein.

15 The term "modification of thermal stability" refers to the change in the amount of thermal energy that is required to cause a given degree of physical change in a target protein that is bound by one or more ligands, relative to the amount of thermal energy that is required to cause the same degree of physical change in the target protein in the absence of any ligand. Modification of thermal stability can be exhibited as an increase or a decrease in thermal stability. Modification of the thermal stability of a protein by a ligand indicates that the ligand binds to the protein. Modification of the thermal stability of a protein by more than one ligand indicates that the ligands bind to the protein.

25 The "midpoint temperature,  $T_m$ " is the temperature midpoint of a thermal unfolding curve. The  $T_m$  can be readily determined using methods well known to those skilled in the art. See, for example, Weber, P. C. *et al.*, *J. Am. Chem. Soc.* 116:2717-2724 (1994); and Clegg, R.M. *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 90:2994-2998 (1993).

30 As discussed above, it is preferable to determine the effect of one or more molecules on the thermal stability of a target protein according to a change in the

$T_m$  of the thermal unfolding curve for the protein. Alternatively the effect of one or more molecules on the thermal stability of a target protein can be determined according to the change in entire thermal unfolding curve for the target protein.

The term "fluorescence probe molecule" refers to an extrinsic fluorophore, which is a fluorescent molecule or a compound which is capable of associating with an unfolded or denatured receptor and, after excitement by light of a defined wavelength, emits fluorescent energy. The term fluorescence probe molecule encompasses all fluorophores. More specifically, for proteins, the term encompasses fluorophores such as thioinosine, and N-ethenoadenosine, formycin, dansyl, dansyl derivatives, fluorescein derivatives, 6-propionyl-2-(dimethylamino)-naphthalene (PRODAN), 2-anilinonaphthalene, and N-aryl-amino-naphthalene sulfonate derivatives such as 1-anilinonaphthalene-8-sulfonate (1,8-ANS), 2-anilinonaphthalene-6-sulfonate (2,6-ANS), 2-amino-naphthalene-6-sulfonate, N,N-dimethyl-2-aminonaphthalene-6-sulfonate, N-phenyl-2-aminonaphthalene-6-sulfonate, N-cyclohexyl-2-aminonaphthalene-6-sulfonate, N-phenyl-2-amino-naphthalene-6-sulfonate, N-phenyl-N-methyl-2-aminonaphthalene-6-sulfonate, N-(o-toluidyl)-2-amino-naphthalene-6-sulfonate, N-(m-toluidyl)-2-amino-naphthalene-6-sulfonate, N-(p-toluidyl)-2-aminonaphthalene-6-sulfonate, 2-(p-toluidinyl)-naphthalene-6-sulfonic acid (2,6-TNS), 4-(dicyanovinyl) julolidine (DCVJ), 6-dodecanoyl-2-dimethylaminonaphthalene (LAURDAN), 6-hexadecanoyl-2-(((2-(trimethylammonium)ethyl)methyl)-amino)naphthalene chloride (PATMAN), Nile red, N-phenyl-1-naphthylamine, 1,1-dicyano-2-[6-(dimethylamino)naphthalen-2-yl]propene (DDNP), 4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid (bis-ANS), and 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole derivative dyes, sold under the trademark DAPOXYL™ (Molecular Probes, Inc., Eugene, OR), including the 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole dyes provided in Diwu, Z. *et al.*, *Photochemistry and Photobiology* 66(4): 424-431 (1997), and in *BioProbes* 25: pp. 8-9, Molecular Probes, Inc., Eugene, OR (1997).

Examples of 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole derivative dyes, and the corresponding Molecular Probes catalogue number, include 5-(4"-

dimethylaminophenyl)-2-(4'-phenyl)oxazole butylsulfonamide (D-12801), 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole-(2-aminoethyl)sulfonamide (D-10460), 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole butylsulfonamide (D-12801), 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole-3-sulfonamidophenylboronic acid (D-10402), 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole sulfonic acid, sodium salt (D-12800), 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole sulfonyl hydrazine (D-10430), 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole-(2-bromoacetamidoethyl)sulfonamide (D-10300), 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole-2-(3-(2-pyridyldithio)propionamidoethyl)sulfonamide (D-10301), 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole sulfonyl chloride (D-10160), 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole-3-sulfonamidopropionic acid, succinimidyl ester (D-10162), 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole carboxylic acid, succinimidyl ester (D-10161).

Preferably the term "fluorescence probe molecule" refers to 1,8-ANS or 2,6-TNS, and 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole derivative dyes, sold under the trademark DAPOXYL™, such as those provided in Diwu, Z. *et al.*, *Photochemistry and Photobiology* 66(4): 424-431 (1997). Still more preferably, the term refers to 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole derivative dyes, sold under the trademark DAPOXYL™, such as those provided in Diwu, Z. *et al.*, *Photochemistry and Photobiology* 66(4): 424-431 (1997). Most preferably, the term refers to 5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole sulfonic acid, sodium salt (D-12800).

The term "carrier" encompasses a platform or other object, of any shape, which itself is capable of supporting at least two containers. The carrier can be made of any material, including, but not limited to glass, plastic, or metal. Preferably, the carrier is a multiwell microplate. The terms microplate and microtiter plate are synonymous. The carrier can be removed from the heating element. In the present invention, a plurality of carriers are used. Each carrier holds a plurality of containers.

The terms "spectral measurement" and "spectrophotometric measurement" are synonymous and refer to measurements of changes in the absorption of light. Turbidity measurements, measurements of visible light absorption, and measurement of ultraviolet light absorption are examples of spectral measurements. Measurement of the intrinsic fluorescence of a target protein, and the fluorescence of an extrinsic fluorophore that is complexed with or bound to a target protein are also examples of spectral measurement and spectrophotometric measurement.

The term "polarimetric measurement" relates to measurements of changes in the polarization properties of light and fluorescent emission. Circular dichroism and optical rotation are examples of polarization properties of light which can be measured polarimetrically. Measurements of circular dichroism and optical rotation are taken using a spectropolarimeter. "Nonpolarimetric" measurements are those that are not obtained using a spectropolarimeter.

Knowledge of the cellular and/or biological function of proteins can be a valuable asset in drug discovery, where it can be useful in developing a detailed understanding of the therapeutic hypothesis for drug function, in designing specific strategies for drug design, and in revealing potential drug side effects.

There are tens of thousands of different enzymes and receptors that constitute potential drug targets, and more are constantly being discovered through genome sequencing studies. These proteins and cellular receptors have specific functions in the biological system, which are practically defined by the molecular ligands with which they form specific interactions. Typical interactions that have functional significance include enzyme interactions with molecular ligands like substrates or substrate analogs, cofactors, adaptor domains, nucleic acids, etc., and receptor interactions with specific ligands, other receptors, cell surface structural components, nucleic acids, polysaccharides, etc.

While it will be broadly possible to isolate, or to clone and express proteins that are putative drug targets, in many cases there will be no functional knowledge base about the protein that can assist in succeeding stages of the drug discovery

process. However, a substantial fraction of known protein molecules fall into mechanistic classes which share important characteristics, including their ability to bind specific types of molecular ligands, including enzyme cofactors, enzyme substrates or substrate analogs, etc. Consequently, it is possible to classify many proteins of otherwise unknown function by their ability to specifically bind various kinds of ligands, either alone or in combination.

When a protein binds to a biological ligand in a functionally significant way, there is an effect on the physical state of the protein that is reflected in its stability relative to its unliganded state. Consequently, one can classify functionally a protein of previously unknown function by incubating it with a probe panel of biological ligands and cofactors (a functional probe library), and measuring which ligands have effects on the stability of the protein. Alternatively, one can determine a previously unknown function of a protein of previously known function by incubating it with a probe panel of biological ligands and cofactors (a functional probe library), and measuring which ligands have effects on the stability of the protein.

As has been established from thermodynamic studies of protein-ligand interactions, when two molecules associate to form a favorable and specific interaction complex, the binding interactions are associated with a reduction in the total free energy of the complex and a net stabilization of the protein-ligand complex relative to the unliganded protein. In practical terms, this means that when an enzyme or receptor interacts with its specific cofactors, or analogs of cofactors, the enzyme or receptor will be stabilized by the interactions. However, it is possible that special situations may exist in which ligand binding may destabilize the target protein. For example, some proteins contain more than one domains or allosteric sites to which one or more ligands can bind.

### ***Overview of the methods of the Present Invention***

The methods of the present invention, as well as other information, are depicted in Figures 1A and 1B.

#### ***A. Identification of a Putative Target Gene.***

5 Target proteins are proteins for which binding to a drug may have therapeutic potential and whose functional characterization may be useful in the drug discovery process. Many genes that are potential targets for therapeutic intervention are identified through a phenomenological correlation that relates a genetic defect to a disease state (e.g., when an inherited disease is correlated with  
10 a genetic defect in a specific enzyme or receptor) or through differences in protein expression patterns in diseased vs. normal tissues.

In many cases it is possible to determine some "function" of a gene product through sequence homology with a homologous protein about which functional or structural data known. However, in a substantial fraction of cases, sequence  
15 homology may not be sufficient to establish functional relationships, and an alternative means is needed to establish function in a way that can directly facilitate the drug discovery process.

#### ***B. Clone and Express the Protein***

To practice the methods of the present invention, it is necessary to obtain  
20 the target protein in sufficient quantities for a biological assay. Proteins that are potential new therapeutic targets and/or require functional characterization may be isolated directly from a natural source using a variety of established biochemical isolation procedures.

The availability of complete gene sequences from genome sequence data  
25 facilitates the cloning and expression of protein targets identified via genomic



methods. For example, the known target DNA sequence may be used to design oligonucleotide probes to select full-length cDNA clones containing the entire cDNA coding for the gene of interest from a representative library of many such cDNA clones. In another example, the known target DNA sequence may be used to design PCR primers for selective amplification and cloning of the gene of interest from total genomic DNA. These and other methods for high-throughput cloning and expression are well-known to those of ordinary skill in the art. Thus, full-length gene sequence data automatically provides the direct means for high-throughput, parallel production of protein targets, an necessary first step in any molecule-based, high-throughput functional screening strategy.

### *C. Thermal Stability Screen*

In order to perform a microplate thermal shift assay of a target protein, is necessary to determine assay conditions that are optimal for carrying out the assay. Proteins are linear polymers of amino acids that spontaneously fold into stable, highly organized 3-dimensional structures. The biological activity and functions of a target protein, including virtually all of the specific binding and catalytic properties that characterize the protein, depends on its three-dimensional structure.

Virtually all folded, active protein domains behave thermally as organic crystals that melt with a cooperative, well defined, pseudo first order phase transition: i.e., melt into a partially disordered, organic liquid-like state, with a well defined melting temperature ( $T_m$ ) that reflects the free energy of stabilization of the protein three-dimensional structure in the experimental solvent conditions. The microplate thermal shift technology uses environment-sensitive fluorescent dyes to sensitively detect the thermal unfolding process and to directly monitor effects on protein stability that arise from perturbations of the solvent environment or through ligand binding to the protein.

5 The stability of the three-dimensional folded state of a protein can potentially be perturbed in several ways. One way is to alter the aqueous solvent environment in which the protein molecules initially fold from a disorganized polymer into the 3-dimensionally organized state. By changing the bulk solvent properties around the protein, the stability of the folded state can be altered relative to the stability of the unfolded state. This can provide a useful strategy for finding optimal conditions for measuring ligand binding and is the principle behind the stability screen.

***D. Microplate Thermal Shift Assay Optimization Screen***

10 The assay optimization screen is a set of solvent conditions and fluorescent dyes that are used with the target protein to determine optimal conditions for performing the microplate thermal shift assay. The protein is subjected to a variety of solution conditions and or fluorescent dyes in order to evaluate the behavior of the protein and/or the assay readout.

15 Examples of variations in conditions could include the addition of organic solvents, variations in pH, salts, etc. that have the potential to alter the relative stability of the folded and unfolded states of the protein. Examples in variations in dyes could include those whose differences in charge, polarity, excitation wavelength, emission wavelength, background signal intensity, or other properties  
20 that offer advantages in precision of measurement, miniaturization or optimization of signal to noise under specific assay conditions. The optimization of conditions that facilitate the stability screen is an empirical process and can readily be practiced by one of ordinary skill in the art.

***E. Functional Probe Library***

25 A substantial fraction of protein molecules that can serve as potential drug targets fall into mechanistic classes which share important characteristics. For

example, many enzymes use ATP as an energetic cofactor, others use pyridine nucleotides as cofactors, some use both as cofactors, etc.

By examining the scientific literature or through experimental means, it is possible to compile a set of enzyme substrates, substrate analogs, cofactors, adaptor protein domains, nucleic acid analogs, polysaccharides, fatty acids, nucleic acids, effector peptides, or other molecules which have been determined to specifically bind to a defined class of protein molecules, or where functional significance has been attached to tight binding to a functionally known class of molecules.

As used herein, a "functional probe library" refers to one or more different molecules that are tested for their ability to bind to a target protein and modify the thermal stability of the protein in response to thermal unfolding. By performing a thermal stability test (preferably by using the microplate thermal shift assay technology) on the protein in the presence of each member the functional probe library, compounds may be incubated with the target protein individually and/or in groups to determine which ligands individually or in combination bind tightly and specifically to the target protein.

Examples of molecules that can comprise a functional probe library include, but are not limited to the following.

### *1. Vitamins and Coenzymes*

NADH/NAD, NADPH/NADP, ATP/ADP, ATP- $\gamma$ -S, acetyl-CoA, biotin, S-adenosyl-methionine, thiamine pyrophosphate (TPP), sulfated oligosaccharides, heparin-like oligosaccharides, GTP, GTP- $\gamma$ -S, gamma-S, pyridoxal-5-phosphate, flavin mononucleotide (FMN), flavin adenine dinucleotide (FAD), folic acid, tetrahydrofolic acid, methotrexate, vitamin K<sub>1</sub>, vitamin E succinate salt, vitamin D<sub>3</sub>, vitamin D<sub>3</sub>-25-hydroxy, vitamin D<sub>3</sub>-1- $\alpha$ -25-dihydroxy, vitamin B<sub>12</sub>, vitamin C, vitamin B<sub>6</sub>, coenzyme A, coenzyme A-n-butyryl, transretinoic acid, and heme.

## 2. *Amino acid residue functional groups and their mimics*

### Building Blocks:

Guanidino groups

Imidazole groups

5 Phenyl groups

Phenolic groups

Indole groups

Aliphatic chains

Single Amino acids and blocked derivatives

10 Higher order structures:

Peptide hormones

Vasopressin

Insulin

TRH

15 Corticotropin

Glucagon

SH<sub>2</sub> domains, SH<sub>3</sub> domains, plextrin  
domains, etc.

Bioactive Peptides

20 Lectins

## 3. *Metal Chelators*

Calcium Chelators (Calbiochem, San Diego, CA)

Iron chelators

## 4. *Metal Ions*

25 Transition metals

Calcium, magnesium

-27-

### 5. *Carbohydrates*

#### Building blocks

Glucose

Galactose

5

Xylose

#### Higher order biomolecules:

Cellulose

Starch

Fructose

10

Mannose

Sucrose

Lactose

Bioactive Carbohydrates (available from Sigma Chemical Co., St. Louis,  
MO)

15

### 6. *Nucleic acids*

#### Building blocks:

Uracil

Thymidine

Cytosine

20

Adenine

Guanine

#### Higher order structures:

Oligonucleotides

Deoxyribonucleic acid (DNA)

25

Ribonucleic acid (RNA)

The methods of the present invention can also be used to screen proteins against libraries of synthetic and naturally-occurring nucleic acids (for example, oligonucleotides) to probe for different classes of nucleic acid-binding proteins. For example, there are many DNA-binding proteins that can be identified by their

ability to bind to particular classes of DNA sequences. Large libraries containing many different nucleic acid sequences (for example, the 4096 different possible synthetic hexamers, can be purchased or synthesized. At high concentrations, all or part of the cognate binding site of site-specific nucleic acid binding proteins can be detected. In the event that a protein appears to bind several different sequences, binding sites can be reconstructed by synthesizing various combinations of nucleic acid sequences, and then the microplate thermal shift assay, or another assay, can be used to measure binding affinities.

There are many DNA-binding proteins that can be identified by their ability to bind to particular classes of DNA sequences with lower specificity. For example, it is well-known that some transcription factors bind a variety of A/T-rich sequences in preference to G/C-rich sequences. Telomerases are known to recognize G/C-rich sequences. Helicases are known to bind short fragments of single-stranded DNA with low specificity. A smaller, more generic library could contain the following components for detecting these and other DNA-binding proteins:

-AT-rich tracts:

$d(T)_{32}/d(A)_{32}$   
 $d(ATAT)_8/d(TATA)_8$   
 $d(AAAT)_8/d(TTTA)_8$   
 $d(AAATT)_6/d(TTTAA)_6$   
 $d(AAATTT)_6/d(TTTAAA)_6$   
 $d(AAAATTTT)_4/d(TTTTAAAA)_4$

-GC-rich tracts:

$d(C)_{32}/d(G)_{32}$   
 $d(GCGC)_8/d(CGCG)_8$   
 $d(GGGCCC)_6/d(CCCGGG)_6$   
 $d(GGGGCCCC)_4/d(CCCCGGGG)_4$

-other

$d(CA)_{32}/d(GT)_{32}$

-29-

 $d(CT)_{32}/d(GA)_{32}$  $d(AG)_{32}/d(TC)_{32}$ 

-Single-stranded components of the above duplex sequences.

5      -d(T)<sub>40</sub>/d(A)<sub>20</sub> (an example of a fragment containing both single-stranded  
and duplex DNA)

-Sheared human chromosomal DNA

- "whole genome amplification" applied to different human chromosomes

-Sheared salmon-sperm DNA

-Sheared microbial DNA

10      -Supercoiled plasmid DNA

-PCR-amplification products from specific chromosomal regions (e.g.  
telomeres and centromeres)

-Other known recognition sites for transcription, RNA processing,  
transposition

15                      7.      *Lipids*

Building blocks:

Choline

Phosphoric acid

Glycerol

20                      Palmitic acid

Oleic acid

Cholesterol

Higher order structures:

Phosphatidyl choline

25                      8.      *Enzyme Inhibitors*

Protease Inhibitors (Sigma Chemical Co., St. Louis, MO)

-30-

PMSF

Leupeptin

Pepstatin A

Bestatin

5                   Peptide aldehyde Cystatin (Cysteine protease inhibitors)

Protein Tyrosine Kinase inhibitors (Calbiochem, San Diego, CA)

Protein Phosphatase inhibitors (Calbiochem, San Diego, CA)

Protein Kinase inhibitors (Calbiochem, San Diego, CA)

Protein Kinase Activators (Calbiochem, San Diego, CA)

10                  Phosphodiesterase inhibitors (Calbiochem, San Diego, CA)

Phospholipase Inhibitors

Transition State Analogs

15                  Similarly, zinc metalloproteases, such as angiotensin converting enzyme, and carboxypeptidase, would be identifiable (a) by destabilization by EDTA or orthophenanthroline ( $Zn^{2+}$  chelation) and (b) by stabilization in the presence of hydroxamates and phosphoramidates that mimic the transition state for  $Zn^{2+}$  catalyzed peptide bond hydrolysis.

The functional probe library can also include steroid compounds amine hormones, and alkaloid compounds.

20                  The functional probe library can be a library of generic drugs. Alternatively, the functional probe library can be a natural product library. For example, see the *Encyclopedia of Common Natural Ingredients Used in Foods, Drugs and Cosmetics*, 2<sup>nd</sup> Edition, Leung and Foster, Eds., Wiley Interscience (1996).

## 25                  F.       Functional Probe Screen

In addition to optimizing conditions that modify protein stability, another way to affect the stability of a folded protein is to specifically bind molecules to



either the unfolded or folded state of the protein. Since virtually all biologically active proteins are folded with organized three-dimensional structures, most interest attaches to ligand molecules that bind to and stabilize the folded state of a protein.

5           As discussed above, a functional probe screen is an assay of the ability of a multiplicity of different molecules in the functional probe library to bind to the target protein and modify the stability of the target protein in response to thermal unfolding. Using the technology, one can directly measure the binding affinity of a small or large molecule ligand to a target protein through its effect on the  
10           unfolding midpoint temperature  $T_m$  (or thermal unfolding profile) of the protein. For molecules that bind to the folded state of the protein, which include most ligands of biological interest, there is a quantitative relationship between the affinity of ligand binding and the extent to which the  $T_m$  of the protein in the liganded state is shifted relative to the  $T_m$  of the protein in unliganded state.

15           Most proteins have functions that are reflected by their ability to bind either large or small molecule ligands with high specificity and high affinity. Many proteins belong to functional classes (e.g. kinases, phosphatases, pyridine nucleotide dependent oxidoreductases, etc.) that bind specific cofactors or catalyze specific reactions using a limited set of catalytic mechanisms.  
20           Consequently, molecules in a given functional class like kinases, which use ATP as a cofactor, will generally bind an non-hydrolyzable ATP cofactor analog like AMPPNP, a property that will be detectable using the methods of the present invention.

25           Moreover, many proteins will bind a combination of ligands or make multiple sets of interactions with biological adaptor domains. To the extent that these interactions are independent, they will generally produce additive perturbations on the stability of the unliganded form of the protein.

30           When the a protein has been tentatively assigned to a particular protein class, one can rescreen the protein using a library of compounds or molecules known to bind to that class of proteins.

### G. *Activity Spectrum*

After performing a thermal stability test (preferably by using the microplate thermal shift assay technology) on the protein in the presence of each member the functional probe library, one can determine which ligands bind tightly and specifically to the target protein and modify the thermal stability of the target protein. The list of compounds (i.e., ligands) that bind to the target protein and modify the thermal stability of the target protein, and the respective affinities of the ligands for the target protein comprise the activity spectrum of the target protein.

### H. *Functional Reference Spectrum List*

As discussed above, a "functional reference spectrum list" is a list of target protein classes (including references to appropriate electronic databases), associated ligands, and corresponding binding constants, that can be used to functionally classify a target protein. Alternatively, the functional reference spectrum list can be a set of one or more activity spectra for one or more known proteins.

As discussed above, a "functional reference list" is a list of proteins that share one or more common features, such as binding to a particular ligand, or exhibiting a common activity. An example of a functional reference list is given in Table 1. The features shared by the proteins listed in Table 1 is that they bind NAD and exhibit dehydrogenase activity. The list of proteins in Table 1 illustrates how a functionally related class of proteins can be discriminated according to their ability to bind different sets of ligands. For example, a protein that binds nicotinamide adenine dinucleotide (NAD), NADPH, or NADH, and malate, as shown by the ability of these compounds to modify the thermal stability of the protein, could be classified as a malate dehydrogenase. As another example, a

protein for which thermal stability is modified by ethanol and NAD could be classified as an alcohol dehydrogenase.

**Table 1**

<b>Functional Reference List</b>	
5	Class 3 Aldehyde Dehydrogenase Human $\delta$ Alcohol Dehydrogenase $\alpha$ -hydroxysteroid Dehydrogenase Malate Dehydrogenase Horse Liver Alcohol Dehydrogenase
10	Alcohol Dehydrogenase Glyceraldehyde-3-Phosphate Dehydrogenase Human $\beta$ -Alcohol Dehydrogenase Dihydropteridine Reductase D-2-Hydroxyisocaproate Dehydrogenase
15	Brassica Napus Enoyl Acp Reductase 7- $\alpha$ -hydroxysteroid Dehydrogenase Holo-D-Glyceraldehyde-3-Phosphate Dehydrogenase Glutathione Reductase D-Glyceraldehyde-3-Phosphate Dehydrogenase
20	Glutathione Reductase 3-Isopropylmalate Dehydrogenase Human $\beta$ -3 Alcohol Dehydrogenase Isocitrate Dehydrogenase Horse Liver Alcohol Dehydrogenase
25	M4 Lactate Dehydrogenase Dihydrolipoamide Dehydrogenase Udp-Gal 4-Epimerase

Table 1

---

**Functional Reference List**

---

D-3-Phosphoglycerate Dehydrogenase  
Human Liver  $\chi\chi$  Alcohol Dehydrogenase  
Alpha, 20  $\beta$ -hydroxysteroid Dehydrogenase  
L-Lactate Dehydrogenase  
NADH Peroxidase

5

***I. Activity Spectrum Comparator***

As used herein, an "activity spectrum comparator" is either a computational or a graphical means by which one can compare the activity spectrum, derived from observing the effects of the functional probe library on the target protein, with the functional reference spectrum list. For example, the activity spectrum comparator can be spreadsheet software that is readily available to those of ordinary skill in the art. For example, MicroSoft Excel (MicroSoft Inc., Redmond, WA) can be used.

10

***J. Functional Classification***

In the methods of the present invention, protein function is indicated by the pattern of ligands that bind to the protein. By using the activity spectrum comparator to compare the observed target activity spectrum with the functional reference spectrum list, the target protein can be functionally classified according to relational data obtained for known proteins. For example, the protein can be classified according to the set of ligands that stabilize the protein against thermal unfolding.

15

20

Thus, by comparing a plot of the degree to which each of a multiplicity of molecules or compounds modify the thermal stability of a protein (and therefore

bind to the protein) to a plot of the degree to which the same molecules modify the thermal stability of a known protein (and therefore bind to the protein), the class of proteins to which the protein belongs can be deduced.

Alternatively, the protein can be classified by comparing the activity spectrum of the target protein to the activity spectra of known, classified proteins. For example, one can consult databases such as PDR online, Medline, SciFinder, STNExpress, in-house databases, NAPRALERT Online, the *Encyclopedia of Common Natural Ingredients Used in Foods, Drugs and Cosmetics*, 2<sup>nd</sup> Edition, Leung and Foster, Eds., Wiley Interscience (1996), and the *Handbook of Enzyme Inhibitors, Part A and B*, 2<sup>nd</sup> Edition, Ellner, Ed., ECH (1990).

#### ***The Microplate Thermal Shift Assay and Apparatus***

In principle, any means of measuring the effects of incubating a protein in the presence of a panel of probe ligands to determine which probe ligands can affect the stability of the target protein will suffice as a means of functionally classifying proteins. Preferably, the microplate thermal shift assay is used to determine the effect of one or more molecules or ligands on the thermal stability of a target protein. The microplate thermal shift assay is a direct and quantitative technology for assaying the effect of one or more molecules on the thermal stability of a target protein.

The thermal shift assay is based on the ligand-dependent change in the thermal unfolding curve of a receptor, such as a protein or a nucleic acid. When heated over a range of temperatures, a receptor will unfold. By plotting the degree of unfolding as a function of temperature, one obtains a thermal unfolding curve for the receptor. A useful point of reference in the thermal unfolding curve is the temperature midpoint ( $T_m$ ), the temperature at which half of the receptor molecules are unfolded.

Thermal shift assays are based on the ligand-dependent change in the midpoint for thermally induced unfolding curves,  $\Delta T_m$ , for the ligand-receptor

complex (relative to the un-complexed receptor) as an experimental observable that directly relates to the ligand binding affinity,  $K_d$ , due to the coupling of the ligand binding and receptor unfolding free energy functions (Schellman, J.A., *Biopolymers* 15: 999-1000 (1976); Brandts, J.F., *Biochemistry* 29:6927-6940 (1990)). This thermal physical screening strategy utilizes the thermal stability of ligand-receptor mixtures as an indicator of the binding affinity for the ligand-receptor interactions. These assays have been traditionally carried out one at a time in differential scanning calorimeters (DSC) that monitor the change in heat capacity as proteins undergo temperature induced unfolding transitions (Brandts *et al.*, *Biochemistry* 29:6927-6940 (1990); and Weber, P. *et al.*, *J. Am. Chem. Soc.* 116:2717-2724 (1994)). Alternatively, thermal shift assays can be performed, again one at a time, by employing temperature-regulated optical instruments that monitor the absorbance (Chavan, A.J. *et al.*, *Biochemistry* 33:7193-7202 (1994)); fluorescence (Chavan, A.J. *et al.*, *Biochemistry* 33:7193-7202 (1994); or circular dichroism (Bouvier, M. *et al.*, *Science* 265:398-402 (1994); Morton, A. *et al.*, *Biochemistry* 34:8564-8575 (1995)) changes that occur for the thermally induced unfolding transitions of proteins.

There are many advantages to using the thermal shift assay since it does not require radioactively labeled compounds, nor fluorescent or other chromophobic labels to assist in monitoring binding. The assay takes advantage of thermal unfolding of biomolecules, a general physical chemical process intrinsic to many, if not all, drug target biomolecules. General applicability is an important aspect of this assay since it obviates the necessity to invent a new assay every time a new therapeutic receptor protein becomes available. The assay is particularly well suited for measuring the binding of ligands to non-enzymatic targets, for example growth factor/receptor interactions, where no spectrophotometric assay is usually possible. However, the single assay configuration of the thermal shift methods, as conventionally performed, has limited the utility of this technique, especially for the high throughput screening of compound libraries.

We have been able to greatly accelerate the protein/ligand screening process by developing a generally applicable high throughput ligand-receptor screening strategy in a 96 well plate (or higher density) format that will identify and rank lead compounds based on the thermodynamic stabilization of ligand-receptor complexes.

Ligand binding stabilizes the receptor (Schellman, J., *Biopolymers* 14:999-1018 (1975)). The extent of binding and the free energy of interaction follow parallel courses as a function of ligand concentration (Schellman, J., *Biophysical Chemistry* 45:273-279 (1993); Barcelo, F. *et al.*, *Chem. Biol. Interactions* 74:315-324 (1990)). As a result of stabilization by ligand, more energy (heat) is required to unfold the receptor. Thus, ligand binding shifts the thermal unfolding curve. That is, ligand binding increases the thermal stability of the protein. This property can be exploited to determine whether a ligand binds to a receptor: a change, or "shift", in the thermal unfolding curve, and thus in the  $T_m$ , suggests that a ligand binds to the receptor.

The thermodynamic basis for the thermal shift assay has been described by Schellman, J.A. (*Biopolymers* 15:999-1000 (1976)), and also by Brandts *et al.* (*Biochemistry* 29:6927-6940 (1990)). Differential scanning calorimetry studies by Brandts *et al.* (*Biochemistry* 29:6927-6940 (1990)) have shown that for tight binding systems of 1:1 stoichiometry, in which there is one unfolding transition, one can estimate the binding affinity at  $T_m$  from the following expression:

$$K_L^{T_m} = \frac{\exp\left\{-\frac{\Delta H_u^{T_0}}{R}\left[\frac{1}{T_m} - \frac{1}{T_0}\right] + \frac{\Delta C_{pu}}{R}\left[\ln\left(\frac{T_m}{T_0}\right) + \frac{T_0}{T_m} - 1\right]\right\}}{L_{T_m}} \quad (\text{equation 1})$$

where  $K_L^{T_m}$  = the ligand association constant at  $T_m$ ;

$T_m$  = the midpoint for the protein unfolding transition in the presence of ligand;

-38-

$T_0$  = the midpoint for the unfolding transition in the absence of ligand;

$\Delta H_u^0$  = the enthalpy of protein unfolding in the absence of ligand at  $T_0$  ;  
 $\Delta C_{pu}$  = the change in heat capacity upon protein unfolding in the absence of  
 ligand;

5

$[L]_m$  = the free ligand concentration at  $T_m$  ; and

$R$  = the gas constant.

This expression was found to be useful for the structure based design of  
 azobenzene ligands for streptavidin where DSC scans of various  
 ligand/streptavidin mixtures facilitated the measurement of binding affinity at  $T_m$   
 (Weber, P. *et al.*, *J. Am. Chem. Soc.* 116:2717-2724 (1994)). These  
 measurements were checked further by performing mixing or isothermal titrating  
 calorimetry experiments which yielded binding affinities consistent with those  
 determined by DSC. The ease and reproducibility of using protein thermal  
 unfolding to estimate ligand binding affinity impressed upon us the potential of  
 further extending this approach for becoming a more general drug discovery tool.

15

20

25

The parameters  $\Delta H_u$  and  $\Delta C_{pu}$  are usually observed from DSC experiments  
 and are specific for each protein. Calorimetric measurements of  $\Delta H_u$  and  $\Delta C_{pu}$  are  
 the most accurate estimates of these parameters because calorimeters typically  
 collect unfolding data every 0.1 °C. However, the parameters,  $\Delta H_u$  and  $\Delta C_{pu}$ , can  
 also be estimated in the microplate thermal shift assay, in which case the  $\Delta H_u$  will  
 not be a calorimetric enthalpy but a comparable van't Hoff enthalpy based on  
 unfolding data collected at every 2.0 °C using the current protocol. Moreover,  
 even in the absence of optimum data for  $\Delta H_u$  and  $\Delta C_{pu}$ , these parameters are  
 constants specific to the protein involved in the compound screening and will  
 therefore be unchanged from well to well, resulting in no influence on calculations  
 of the relative values of binding affinities, i.e.,  $K_L$  at  $T_m$ .

Besides the parameters  $\Delta H_u$  and  $\Delta C_{pu}$ , it is also necessary to obtain  
 estimates of  $T_m$  and  $T_0$  to solve for  $\kappa_L^T$  in equation 1. This is accomplished



through the use of non-linear least squares computer fits of the unfolding data for each individual well using the following equation:

*Equation 2*

$$y(T) = y_u + \frac{(y_f - y_u)}{1 + \exp\left(\left[\frac{-\Delta H_u}{R}\right]\left[\frac{1}{T} - \frac{1}{T_m}\right] + \left[\frac{\Delta C_{pu}}{R}\right]\left[\left(\frac{T_m}{T} - 1\right) + \ln\left(\frac{T}{T_m}\right)\right]\right)}$$

Equation 2 employs five fitting parameters,  $\Delta H_u$ ,  $\Delta C_{pu}$ ,  $T_m$ ,  $y_f$  and  $y_u$ , where  $y_f$  and  $y_u$  are the pre-transitional and post-transitional fluorescence levels, respectively. The computer fits are determined by floating these parameters to arrive at the minimum of the sum of the squares of the residuals by employing the Levenberg-Marquardt algorithm. The  $T_0$  values are obtained for wells that contain no added ligand and are set as the reference. Commercially available curve-fitting software is readily available to one of ordinary skill in the art. For example, Kaleidograph 3.0 (Synergy, Reading, PA) can be used.

It is also possible to calculate the ligand association equilibrium constant at any temperature,  $K_L$  at  $T$ , the ligand association equilibrium constant at  $T_m$ , using equation 3, if mixing calorimetry data for the binding enthalpy at  $T$ ,  $\Delta H_L$ , and the change in heat capacity upon ligand binding,  $\Delta C_{pL}$ , are known (Brandts & Lin, 1990).

*Equation 3*

$$K_L^T = K_L^{T_m} \exp\left\{-\frac{\Delta H_L^T}{R}\left[\frac{1}{T} - \frac{1}{T_m}\right] + \frac{\Delta C_{pL}}{R}\left[\ln\left(\frac{T}{T_m}\right) - \frac{T}{T_m} + 1\right]\right\}$$

where  $K$  = the ligand association constant at any temperature,  $T$ .

$K_L^{T_m}$  = the ligand association constant at  $T_m$ .

-40-

$T_m$  = the midpoint for the protein unfolding transition in the presence of ligand.

$\Delta H_L^T$  = the enthalpy of ligand binding at temperature,  $T$ .

$\Delta C_p$  = the change in heat capacity upon binding of ligand.

$R$  = gas constant

5           The second exponential term of equation 3 is usually small enough to be ignored so that approximate values of  $K_L$  at  $T$  can be obtained using just the first exponential term, and equation 3 reduces to equation 4:

**Equation 4**

$$K_L^T = K_L^{T_m} \exp \left[ -\frac{\Delta H_L^T}{R} \left[ \frac{1}{T} - \frac{1}{T_m} \right] \right]$$

10           The parameter  $\Delta H_L^T$  can be measured using a isothermal titrating calorimetry, using a calorimetric device such as the Omega (MicroCal; Northampton, MA). When calorimetric data are not available,  $\Delta H_L^T$  can be estimated to be about -10.0 kcal/mol, which is an average binding enthalpy (Wiseman *et al.*, *Anal. Biochem.* 179:131-137 (1989)).

15           Preferably, fluorescence spectrometry is used to monitor thermal unfolding. The fluorescence methodology is more sensitive than the absorption methodology. The use of intrinsic protein fluorescence and fluorescence probe molecules in fluorescence spectroscopy experiments is well known to those skilled in the art. See, for example, Bashford, C.L. *et al.*, *Spectrophotometry and Spectrofluorometry: A Practical Approach*, IRL Press Ltd., pub., pp. 91-114 (1987); Bell, J.E., *Spectroscopy in Biochemistry, Vol. I*, CRC Press, pub., pp. 155-194 (1981); Brandts, L. *et al.*, *Ann. Rev. Biochem.* 41:843 (1972).

20           The microplate thermal shift assay is further described in U.S. patent Appl. No. 08/853,464, filed May 9, 1997, and in international patent Appl. No. PCT/US97/08154 (published November 13, 1997 as publication no. WO 97/42500), which are hereby incorporated by reference in their entirety.

25

Spectral readings, preferably fluorescence readings, can be taken on all of the samples on a carrier simultaneously. Alternatively, readings can be taken on samples in groups of at least two at a time.

5 A fluorescence imaging system, for example, a fluorescence emission imaging system, can be used to monitor the thermal unfolding of a target molecule or a receptor. Fluorescence imaging systems are well known to those skilled in the art. For example, the ALPHAIMAGER™ Gel Documentation and Analysis System (Alpha Innotech, San Leandro, CA) employs a high performance charge coupled device (CCD) camera with 768 x 494 pixel resolution. The charge  
10 coupled device camera is interfaced with a computer and images are analyzed with Image analysis software™. The CHEMIIMAGER™ (Alpha Innotech) is a cooled charge coupled device that performs all of the functions of the ALPHAIMAGER™ and in addition captures images of chemiluminescent samples and other low intensity samples. The CHEMIIMAGER™ charge coupled device  
15 includes a Pentium processor (1.2 Gb hard drive, 16 Mb RAM), AlphaEase™ analysis software, a light tight cabinet, and a UV and white light trans-illuminator. For example, the MRC-1024 UV/Visible Laser Confocal Imaging System (BioRad, Richmond, CA) facilitates the simultaneous imaging of more than one fluorophore across a wide range of illumination wavelengths (350 to 700 nm).  
20 The Gel Doc 1000 Fluorescent Gel Documentation System (BioRad, Richmond, CA) can clearly display sample areas as large as 20 x 20 cm, or as small as 5 x 4 cm. At least two 96 well microplates can fit into a 20 x 20 cm area. The Gel Doc 1000 system also facilitates the performance of time-based experiments.

A fluorescence imaging system, for example, a fluorescence emission  
25 imaging system, can be used to monitor receptor unfolding in a microplate thermal shift assay. In this embodiment, a plurality of samples is heated simultaneously between 25 to 110°C. A fluorescence emission reading is taken for each of the plurality of samples simultaneously. For example, the fluorescence in each well of a 96 or a 384 well microplate can be monitored simultaneously. Alternatively,  
30 fluorescence readings can be taken continuously and simultaneously for each

sample. At lower temperatures, all samples display a low level of fluorescence. As the temperature is increased, the fluorescence in each sample increases. Wells which contain ligands which bind to the target molecule with high affinity shift the thermal unfolding curve to higher temperatures. As a result, wells which contain  
5 ligands which bind to the target molecule with high affinity fluoresce less, at a given temperature above the  $T_m$  of the target molecule in the absence of any ligands, than wells which do not contain high-affinity ligands. If the samples are heated in incremental steps, the fluorescence of all of the plurality of samples is simultaneously imaged at each heating step. If the samples are heated  
10 continuously, the fluorescent emission of all of the plurality of samples is simultaneously imaged during heating.

A thermal shift assay can be performed in a volume of 100  $\mu$ L volumes. For the following reasons, however, it is preferable to perform a thermal shift assay in a volume of 1-10  $\mu$ L. First, approximately 10- to 100-fold less protein  
15 is required for the miniaturized assay. Thus, only ~ 4 to 40 pmole of protein are required (0.1  $\mu$ g to 1.0  $\mu$ g for a 25 kDa protein) for the assay (i.e. 1 to 10  $\mu$ L working volume with a target molecule concentration of about 1 to about 4  $\mu$ M). Thus, 1 mg of protein can be used to conduct 1,000 to 10,000 assays in the miniaturized format. This is particularly advantageous when the target molecule  
20 is available in minute quantities.

Second, approximately 10- to 100-fold less ligand is required for the miniaturized assay. This advantage is very important to researchers when screening valuable combinatorial libraries for which library compounds are synthesized in minute quantities. In the case of human  $\alpha$ -thrombin, the ideal ligand  
25 concentration is about 50  $\mu$ M, which translates into 25-250 pmoles of ligand, or 10-100 ng (assuming a MW of 500 Da) of ligand per assay in the miniaturized format.

Third, the smaller working volume allows the potential of using larger arrays of assays because the miniaturized assay can fit into a much smaller area.  
30 For example, 384 well (16 x 24 array) or 864 well (24 x 36 array) plates have the

same dimensions as the 96 well plates (8.5 x 12.5 cm). The 384 well plate and the 864 well plate allows the user to perform 4 and 9 times as many assays, respectively, as can be performed using a 96 well plate. Alternatively, plates with more wells, such as 1536 well plates (32 x 48 arrays; Matrix Technologies Corp.), can be used. A 1536 well plate will facilitate sixteen times the throughput afforded by a 96 well plate.

Thus, using the 1536 well plate configuration, assay speed can be increased by about 16 times, relative to the speed at which the assay can be performed using the 96 well format. The 8 x 12 assay array arrangement (in a 96-well plate) facilitates the performance of 96 assays/hr, or about 2300 assays/24 hours. The 32 x 48 array assay arrangement facilitates the performance of about 1536 assays hr., or about 37,000 assays/24 hours can be performed using a 32 x 48 assay array configuration.

The assay volume can be 1-100  $\mu$ L. Preferably, the assay volume is 1-50  $\mu$ L. More preferably, the assay volume is 1-25  $\mu$ L. More preferably still, the assay volume is 1-10  $\mu$ L. More preferably still, the assay volume is 1-5  $\mu$ L. More preferably still, the assay volume is 5  $\mu$ L. Most preferably, the assay volume is 1  $\mu$ L or 2  $\mu$ L.

Alternatively, the assay is performed in V-bottom polycarbonate, polystyrene, or polypropylene plates or dimple plates. A dimple plate is a plate that contains a plurality of round-bottom wells that hold a total volume of 15  $\mu$ L.

The microplate thermal shift assay is performed by (a) contacting a protein with one or more of a multiplicity of different molecules in each of a multiplicity of containers; (b) heating the multiplicity of containers from step(a), preferably simultaneously; (c) measuring in each of the containers a physical change associated with the thermal unfolding of the target molecule resulting from heating; (d) generating a thermal unfolding curve for the target molecule as a function of temperature for each of the containers; and (e) comparing each of the unfolding curves in step (d) to (1) each of the other thermal unfolding curves and to (2) the thermal unfolding curve obtained for the protein in the absence of any

of the multiplicity of different molecules; and (f) determining whether any of the multiplicity of different molecules modifies the thermal stability of the protein, wherein a modification in thermal stability is indicated by a shift in the thermal unfolding curve.

5           Step (d) may further comprise determining a midpoint temperature ( $T_m$ ) from the thermal unfolding curve. Step (e) may further comprise comparing the  $T_m$  of each of the unfolding curves in step (d) to (1) the  $T_m$  of each of the other thermal unfolding curves and to (2) the  $T_m$  of the thermal unfolding curve obtained for the target protein in the absence of any of the different molecules.

10           To practice the methods of the present invention using fluorescence spectroscopy or imaging, step (a) comprises contacting the target protein with a fluorescence probe molecule present in each of the multiplicity of containers and step (c) comprises (c1): exciting the fluorescence probe molecule, in each of the multiplicity of containers, with light; and (c2) measuring the fluorescence from  
15 each of the multiplicity of containers. Fluorescence, for example, fluorescence emission, can be measured from each of the multiplicity of containers one container at a time, from a subset of the multiplicity of containers simultaneously, or from each of the multiplicity of containers simultaneously.

20           To generate an activity spectrum, molecules are ranked according to the degree to which they stabilize the target protein against thermal unfolding. After the molecules are ranked, the activity spectrum of the target protein for the molecules in the functional probe library is compared to one or more functional reference spectrum lists.

25           Suitable heating apparatuses for practicing the methods of the present invention are well known to those of ordinary skill in the art. For example, the ROBOCYCLER® Gradient Temperature Cycler (Stratagene, La Jolla, CA) (see U.S. Patent No. 5,525,300) can be used. Alternatively, a temperature gradient heat block can be used (see U.S. Patent No. 5,255,976). Fluorescence can be read using any suitable fluorescence spectroscopy device. For example, the CytoFluor  
30 II apparatus (PerSeptive Biosystems, Framingham, MA) can be used.

The element upon which the sample carrier is heated can be any element capable of heating samples rapidly and in a reproducible fashion. In the present invention, a plurality of samples is heated simultaneously. The plurality of samples can be heated on a single heating element. Alternatively, the plurality of samples  
5 can be heated to a given temperature on one heating element, and then moved to another heating element for heating to another temperature. Heating can be accomplished in regular or irregular intervals. To generate a smooth unfolding curve, the samples should be heated evenly, in intervals of 1 or 2°C. The temperature range across which the samples can be heated is from 4 to 110°C.  
10 Spectral readings, and particularly fluorescence readings, are taken after each heating step. Samples can be heated and read by the spectral device, for example, a fluorescence imaging camera, in a continuous fashion. Alternatively, after each heating step, the samples may be cooled to a lower temperature prior to taking the spectral readings. Preferably, the samples are heated continuously and spectral  
15 readings are taken while the samples are being heated.

Spectral, e.g., fluorescence, readings can be taken on all of the samples in the carrier simultaneously. Alternatively, readings can be taken on samples in groups of at least two at a time. Finally, the readings can be taken one sample at a time.

20 Preferably, the instrument used to perform the microplate thermal shift assay consists of a scanner and a control software system. Fluorescence, for example, fluorescence emission, can be detected by a photomultiplier tube in a light-proof detection chamber. The software runs on a personal computer and the action of the scanner is controlled through the software.

25 An exemplary apparatus 200 is shown in Figure 2. A precision X-Y mechanism scans the microplate with a sensitive fiber-optic probe to quantify the fluorescence in each well. The microplate and samples can remain stationary during the scanning of each row of the samples, and the fiber-optic probe is then moved to the next row. Alternatively, the microplate and samples can be moved  
30 to position a new row of samples under the fiber-optic probe. The scanning

system is capable of scanning 96 samples in under one minute. The scanner is capable of holding a plurality of excitation filters and a plurality of emission filters to measure the most common fluorophores. Thus, fluorescence emission readings can be taken one sample at a time, or on a subset of samples simultaneously.

5           The heat conducting element or block upon which the sample carrier is heated can be any element capable of heating samples rapidly and reproducibly. The plurality of samples can be heated on a single heating element. Alternatively, the plurality of samples can be heated to a given temperature on one heating element, and then moved to another heating element for heating to another  
10           temperature. Heating can be accomplished in regular or irregular intervals. To generate a smooth unfolding curve, the samples should be heated evenly, in intervals of 1 or 2°C. The temperature range across which the samples can be heated is from 4 to 110°C.

          Preferably, a plurality of samples is heated simultaneously. If samples are  
15           heated in discrete temperature intervals, in a stairstep fashion, spectral readings are taken after each heating step. Alternatively, after each heating step, the samples may be cooled to a lower temperature prior to taking the spectral readings. Alternatively, samples can be heated in a continuous fashion and spectral readings are taken during heating.

20           The assay apparatus can be configured so that it contains a single heat conducting block. Alternatively, the assay apparatus can be configured so that it contains a plurality of heat conducting blocks upon a movable platform. The platform may be a translatable platform that can be translated, for example, by a servo driven linear slide device. An exemplary linear slide device is model SA  
25           A5M400 (IAI America, Torrance, CA). In this embodiment, the sensor receives spectral emissions from each of the samples on a given heat conducting block. The platform is then translated to place another heat conducting block and its accompanying samples under the sensor so that it receives spectral emissions from each of the samples on that heating block. The platform is translated until spectral  
30           emissions are received from the samples on all heat conducting blocks



Alternatively, the platform may be a rotatable platform, as shown in Figure 2, that may be rotated, for example, by a servo driven axle. In the latter embodiment, the sensor receives spectral emissions from each of the samples on a given heat conducting block. The platform is then rotated to place another heat conducting block and its accompanying samples under the sensor so that it receives spectral emissions from each of the samples on that heating block. The platform is rotated until spectral emissions are received from the samples on all heat conducting blocks.

In apparatus 200, a plurality of heat conducting blocks 204, each of which includes a plurality of wells for a plurality of samples 210, is mounted on a rotatable platform or carousel 206. Platform or carousel 206 can be composed of a heat conducting material, such as the material that heat conducting block 204 is composed of. Axle 208 is rotatably connected to base 202. Rotatable platform 206 is axially mounted to rotate about axle 208. Rotation of axle 208 is controlled by a servo controller 210. Servo controller 210 is controlled by a computer controller 250 in a manner well known to one of skill in the relevant arts. Computer controller 250 causes servo controller 210 to rotate axle 208 thereby rotating rotatable platform 206. In this manner, heat conducting blocks 204 are sequentially placed under fiber optic probe 212.

Each of the plurality of heat conducting blocks 204 can be controlled independently by temperature controller 214. Thus, the temperature of a first heat conducting block 204 can be higher or lower than the temperature of a second heat conducting block 204. Similarly, the temperature of a third heat conducting block 204 can be higher or lower than the temperature of either first or second heat conducting block 204.

Temperature controller 214 is connected to heat conducting block 204 by a thermoelectric connection 230. Under the action of temperature controller 214, the temperature of heat conducting block 204 can be increased, decreased, or held constant. Temperature controller 214 can be configured to adjust the temperature of rotatable platform 206. In such a configuration, when rotatable

platform 206 is heated, heat conducting blocks 204 are also heated. Alternatively, the temperature of each of heat conducting blocks 204 can be controlled by a circulating water system such as that noted above. Particularly, the temperature of heat conducting block 204 can be changed by temperature controller 214 in accordance with a pre-determined temperature profile. Preferably, temperature computer controller 214 is implemented using a computer system.

As used herein, the term "temperature profile" refers to a change in temperature over time. The term "temperature profile" encompasses continuous upward or downward changes in temperature, both linear and non-linear changes. The term also encompasses any stepwise temperature change protocols, including protocols characterized by incremental increases or decreases in temperature during which temperature increases or decreases are interrupted by periods during which temperature is maintained constant. In the apparatus shown in Figure 2, the temperature profile can be pre-determined by programming temperature computer controller 214. For example, temperature profiles can be stored in a memory device of temperature controller 214, or input to temperature controller 214 by an operator.

Assay apparatus 200 also includes a light source 218 for emitting an excitatory wavelength of light. Excitatory light from light source 218 excites samples 216 with excitatory light. Any suitable light source can be used. Excitatory light causes a spectral emission from samples 216. The spectral emission can be electromagnetic radiation of any wavelength in the electromagnetic spectrum. Preferably, the spectral emission is fluorescent, ultraviolet, or visible light. Most preferably, the spectral emission is fluorescence emission.

A sensor is removably attached to a sensor armature 226. An exemplary sensor is a fiber optic probe 212. Fiber optic probe 212 includes a fiber optic cable capable of transmitting excitatory light to samples 216, and a fiber optic cable capable of receiving a spectral emission from samples 216. Electromagnetic

radiation is transmitted from excitatory light source 218 to fiber optic probe 212 by excitatory light input fiber optic cable 228.

An excitatory light filter servo controller 258 controls the aperture of excitatory light filter 256. Excitatory light source 218 and excitatory light filter servo controller 258 are communicatively and operatively connected to excitatory light computer controller 254. Computer controller 254 controls the wavelength of excitatory light transmitted to samples 216 by controlling excitatory light filter servo controller 258. Excitatory light is transmitted through excitatory light input fiber optic cable 228 to fiber optic probe 212 for transmission to samples 216.

The spectral emission from samples 216 is received by fiber optic probe 212 and is transmitted to a spectral emission filter 238 by output fiber optic cable 250. A spectral emission servo controller 240 controls the aperture of spectral emission filter 238, thereby controlling the wavelength of the spectral emission that is transmitted to photomultiplier tube 220. Spectral emission servo controller 240 is controlled by a computer controller 242.

The spectral emission from samples 216 is transmitted from photomultiplier tube 220. Electrical output 244 connects photomultiplier tube 220 to electric connection 224. Electric connection 224 connects electrical output 244 to computer 222. Driven by suitable software, computer 222 processes the spectral emission signal from samples 216. Exemplary software is a graphical interface that automatically analyzes fluorescence data obtained from samples 216. Such software is well known to those of ordinary skill in the art. For example, the CytoFluor™II fluorescence multi-well plate reader (PerSeptive Biosystems, Framingham, MA) utilizes the Cytocalc™ Data Analysis System (PerSeptive Biosystems, Framingham, MA). Other suitable software includes, MicroSoft Excel or any comparable software.

A sensor armature relative movement means 260 moves sensor armature 226 in directions 234 and 236. A second relative movement means 232 moves sensor armature 226 in directions 246 and 248 so that fiber optic probe 212 can be moved to detect spectral emissions from samples 216.

As discussed above, the spectral receiving means or sensor of the assay apparatus of the present invention can comprise a photomultiplier tube. Alternatively, the spectral receiving means or sensor can include a charge coupled device (CCD). In still another alternative, the spectral receiving means or sensor  
5 can include a diode array. A CCD is made of semi-conducting silicon. When photons of light fall on it, free electrons are released.

Further, a CCD camera can be used to image fluorescence, such as fluorescence emission. High resolution CCD cameras can detect very small amounts of electromagnetic energy, whether it originates from distance stars, is  
10 diffracted by crystals, or is emitted by fluorophores. As an electronic imaging device, a CCD camera is particularly suitable for fluorescence emission imaging because it can detect very faint objects, affords sensitive detection over a broad spectrum range, affords low levels of electromagnetic noise, and detects signals over a wide dynamic range--that is, a charge coupled device can simultaneously  
15 detect bright objects and faint objects. Further, the output is linear so that the amount of electrons collected is directly proportional to the number of photons received. This means that the image brightness is a measure of the real brightness of the object, a property not afforded by, for example, photographic emulsions. Suitable CCD cameras are available from Alpha-Innotech (San Leandro, CA),  
20 Stratagene (La Jolla, CA), and BioRad (Richmond, CA).

Apparatuses useful for practicing the microplate thermal shift assay are further described in U.S. patent Appl. No. 08/853,459, filed May 9, 1997, and in international patent Appl. No. PCT/US97/08154 (published November 13, 1997 as publication no. WO 97/42500), which are hereby incorporated by reference in  
25 their entirety.

Having now generally described the invention, the same will become more readily understood by reference to the following specific examples which are included herein for purposes of illustration only and are not intended to be limiting unless otherwise specified.

**Example 1****Wide Cross Target Utility of Microplate Thermal Shift Assay**

A number of different therapeutic protein targets have been tested in the microplate thermal shift assay, to date, and are listed in Table 2. They include a variety of different proteins, with a wide diversity of *in vivo* function. Included here are various serine proteases, a DNA binding protein (*lac* repressor), two growth factors (basic fibroblast growth factor (bFGF) and acidic fibroblast growth factor (aFGF)), and a growth factor receptor (domain II of the fibroblast growth factor receptor 1 (D(II)FGFR1)).

**Table 2**

<b>Therapeutic Targets Analyzed by the Microplate Thermal Shift Assay</b>			
<b>Targets</b>	<b>MW</b>	<b>Assays/mg (in 10 uL format)</b>	
$\alpha$ -Thrombin	37.0 kDa	1430	0.7 ug/assay (20 pmol)
Factor D	25.0 kDa	1000	1.0 ug/assay (40 pmol)
Factor Xa	45.0 kDa	1667	0.6 ug/assay (7 pmol)
bFGF	17.5 kDa	2000	0.5 ug/assay (29 pmol)
D(II)FGFR1	13.5 kDa	588	1.7 ug/assay 126 pmol)
<i>lac</i> Repressor	77.0 kDa	1200	0.8 ug/assay (10 pmol)
Urokinase	28.0 kDa	714	1.4 ug/assay (50 pmol)
NFkB protein	65.0 kDa	3030	0.33 ug/assay (5 pmol)
GLP1 receptor	26.0 kDa		
MHC II	45.0 kDa		
von Willebrand Factor	500 kDa	400	2.5 ug/assay
aFGF	18.0 kDa		

The molecular weights of the target proteins range from 13.5 kDa to about 500 kDa. On average, it was possible to conduct 1322 assays per 1.0 mg of protein using a 10  $\mu$ L assay volume. The number of assays that can be conducted can be doubled if the 5  $\mu$ L assay format is employed.

5 All microplate thermal shift assays were performed in polycarbonate V-bottom 96 well plates using 200  $\mu$ M 1,8-ANS as the fluorescent probe for monitoring the thermal unfolding transitions for the protein/ligand mixtures. Changes in fluorescence emission at 460 nm were monitored with a CytoFluor II (PerSeptive Biosystems) fluorescence plate reader (excitation at 360 nm), and the  
10 temperature was raised in 2°C increments with the RoboCycler® Gradient Temperature Cycler (Stratagene, La Jolla, CA).

A number of other proteins have been assayed using the microplate thermal shift assay, including the following proteins from the following classes: serine proteases (thrombin, Factor Xa, Factor D, urokinase, trypsin, chymotrypsin, subtilisin); cell surface receptors (FGF receptor 1, MHC Class II, GLP1 receptor,  $\beta$ -2 adrenergic receptor, fibronectin receptor (libIIIa)); growth factors (aFGF, bFGF); DNA binding proteins (*lac* repressor, NF- $\kappa$ -B, helicase); motor proteins (myosin, helicase); oxido-reductases (horseradish peroxidase, cytochrome c, lactate dehydrogenase, lactoperoxidase, malate dehydrogenase, cholesterol  
20 oxidase, glyceraldehyde 3-phosphate dehydrogenase, phosphoenolpyruvate carboxylase, dihydrofolate reductase); carbohydrate modifications (cellulase,  $\alpha$ -amylase, hyaluronidase,  $\beta$ -glucosidase, invertase); immunoglobulins (IgG Fab, IgG Fc); DNases (DNase I, DNase II) RNases (RNase A); intracellular calcium receptors (calmodulin, S100 protein); neurotransmitter hydrolase  
25 (acetylcholinesterase); free radical scavenger (superoxide dismutase); biotin binding protein (streptavidin); oxygen binding protein (myoglobin); and protease inhibitor (trypsin inhibitor).

## ***Example 2***

### ***Multi-Ligand Binding Interactions With A Single Target Protein***

The near universal utility of the microplate thermal shift assay technology is also illustrated for multi-ligand binding interactions that many times occur within a single protein molecule. The ability to assess the binding of many different kinds of ligands to a single protein without re-tooling the assay is a great advantage of this technology and easily lends itself to the task of assigning function to a protein for which nothing is known other than the primary sequence. Knowledge for the binding of different ligands will help in the evaluation of the function of a sample protein derived from genomics information.

As previously demonstrated, the microplate thermal shift assay can be used to screen ligands for binding to single sites on target proteins. However, based upon the near additivity of the free energy of ligand binding and protein unfolding, it is also possible to employ the microplate thermal shift assay to analyzing multi-ligand binding interactions on a target protein. In principle, if the free energy of binding of different ligands binding to the same protein are nearly additive then one can analyze multi-ligand binding systems either non-cooperative or cooperative (positive or negative).

In this regard, human thrombin is an ideal system to test the utility of the assay for analysis of multi-ligand binding interactions because it has at least four different binding sites: (1) the catalytic binding site; (2) the fibrin binding site (exosite I); (3) the heparin binding site (exosite II); (4) the Na<sup>+</sup> binding site, located about 15Å from the catalytic site.

First, the binding of individual ligands was determined. 3DP-4660, Hirugen (hirudin 53-64) (Sigma), and heparin 5000 (CalBiochem), bind to the catalytic site, the fibrin binding site and the heparin binding site, respectively of thrombin.

A stock thrombin solution was diluted to 1 μM in 50 mM Hepes, pH 7.5, 0.1 M NaCl, 1 mM CaCl<sub>2</sub>, and 100 μM 1,8-ANS. Each thrombin ligand was included singly and in various combinations to 1 μM thrombin solutions at final

concentrations of 50  $\mu$ M each, except for heparin 5000, which was 200  $\mu$ M. 100  $\mu$ L of thrombin or thrombin/ligand(s) solution was dispensed into wells of a 96-well V-bottom polycarbonate microtiter plate. The contents were mixed by repeated uptake and discharge in a 100  $\mu$ L pipette tip. Finally, one drop of mineral oil (Sigma, St. Lois, MO) was added on top of each reaction well to reduce evaporation from samples at elevated temperatures. The plate was subjected to 3 minutes of heating in a RoboCycler<sup>®</sup> Gradient Temperature Cycler (Stratagene, La Jolla, CA) thermal block, with which a temperature gradient was created across the microplate, followed by 30 seconds cooling at 25°C, and subsequent reading in a fluorescence plate reader. Data were analyzed by non-linear least squares fitting.

The results of these individual binding reactions are shown in Figure 3. The rank order of binding affinity was 3DP-4660 > hirugen > heparin 5000, corresponding to  $K_d$  of 15 nM, 185 nM and 3434 nM, respectively, for the ligands binding at each  $T_m$  (see Equation (1)).

Next, the binding of combinations of two ligands was studied. The data are shown in Figure 4. The results in Figure 4 reveal thermal unfolding shifts that are slightly smaller than that expected for full additivity. For example, Hirugen alone gave a  $\Delta T_m$  of 5.8°C, and 3DP-4660 alone gave a  $\Delta T_m$  of 7.7°C, but together they gave a  $\Delta T_m$  of 12.2°C, and not the 13.5°C shift that would be expected if the binding energies were fully additive. This result could mean that the binding affinity of one or both ligands is diminished when both ligands are bound to thrombin, and would be an example of negative cooperativity between the fibrin binding site and the catalytic binding site. Such a result is consistent with the thrombin literature, in which the kinetics of hydrolysis of various chromogenic substrates has been found to depend upon ligands binding to exosite I. Indeed, a 60% decrease in  $K_m$  for the hydrolysis of D-phenylalanylpipicolyl arginyl-p-nitroanilide was observed when Hirugen was present (Dennis *et al.*, *Eur. J. Biochem.* 188:61-66 (1990)). Moreover, there is also structural evidence for cooperativity between the catalytic site and exosite I. A comparison of the



isomorphous structures of PPACK-bound thrombin (PPACK is a thrombin catalytic site inhibitor) and hirugen-bound thrombin revealed conformational changes that occur at the active site as a result of hirugen binding at the exosite I (Vijayalakshmi *et al.*, *Protein Science* 3:2254-2271 (1994)). Thus, the apparent cooperativity observed by between the catalytic center and the exosite I are consistent with functional and structural data in the literature.

One would expect that if the energies of binding of all three ligands were fully additive, a  $\Delta T_m$  of 17.7°C would be seen. However, when all three ligands were present together, the  $\Delta T_m$  was 12.9°C. This result implies further negative cooperativity that involves ligand binding at all three protein binding sites. There is some evidence in the literature that is consistent with this supposition. For example, thrombin, in a ternary complex with heparin and fibrin monomer, has decreased activity toward tri-peptide chromogenic substrates and pro-thrombin (Hogg & Jackson, *J. Biol. Chem.* 265:248-255 (1990)), and markedly reduced reactivity with anti-thrombin (Hogg & Jackson, *Proc. Natl. Acad. Sci. USA* 86:3619-3623 (1989)). Also, recent observations by Hotchkiss *et al.* (*Blood* 84:498-503 (1994)) indicate that ternary complexes also form in plasma and markedly compromise heparin anticoagulant activity.

A summary of the thrombin multi-ligand binding results is shown in Table 3. From the results in Figures 3 and 4 and in Table 3, the following conclusions were made. First, in the presence of heparin 5000, hirudin 53-65 bound thrombin about 21 times less tightly than in the absence of heparin; and in the presence of heparin 5000, 3DP-4660 bound thrombin about 10 times less tightly than in the absence of heparin.

Second, in the presence of hirudin 53-65, heparin bound thrombin about 18 times less tightly than in the absence of hirudin 53-65; and in the presence of hirudin 53-65, 3DP-4660 bound thrombin about 3 times less tightly than in the absence of hirudin 53-65.

Third, in the presence of 3DP-4660, heparin bound thrombin about 25% more tightly than in the absence of 3DP-4660; and in the presence of 3DP-4660,

hirudin bound thrombin about 2.3 times less tightly than in the absence of 3DP-4660.

Table 3

Assay for Ligands Binding to the Active Site, Exosite, and Heparin binding Site of Thrombin					
Protein/Ligand	[Ligand] ( $\mu\text{M}$ )	$T_m$ ( $^{\circ}\text{K}$ )	$\Delta T_m$ ( $^{\circ}\text{K}$ )	$K_d$ at $T_m^a$ (nM)	$K_d$ at $298^{\circ}\text{K}^b$ (nM)
Thrombin (TH)	none	323.75	0.0		
TH/Heparin 5000	200	327.95	4.2	3434	470
TH/Hirudin 53-65	50	329.52	5.8	185	23
TH/3dp-4660	50	331.40	7.7	29	3
TH/Heparin 5000	200	327.95			
TH/Hep./Hir.	50	330.57	2.6	4254	478
TH/Heparin 5000	200	327.95			
TH/Hep./3dp 4660	50	333.20	5.3	350	32
TH/Hirudin 53-65	50	329.52			
TH/Hir./Hep.	200	330.57	1.1	75422	8467
TH/Hirudin 53-65	50	329.52			
TH/Hir./3dp-4660	50	335.97	6.5	117	9
TH/3dp-4660	50	331.40			
TH/3dp-4660/Hep.	200	333.20	1.8	38205	351
TH/dp-4660	50	331.40			
TH/3dp-4660/Hir.	50	335.97	4.6	731	54

a: Calculations for  $K_d$  at  $T_m$  were made using equation (1) with  $\Delta H_{\text{ref}}^{\text{ref}}$  = 200.0 kcal/mole, as observed for pre-thrombin 1 by Leintz *et al.*, *Biochemistry* 33:5460-5468 (1994), and an estimated  $\Delta C_p^{\text{ref}}$  = 2.0 kcal/mole  $^{\circ}\text{K}$ ; and  $K_d = 1/K_a$ .

b: Estimates for  $K_d$  at  $T = 298^{\circ}\text{K}$  were made using the equation (3), where  $\Delta H_{\text{ref}}^{\text{ref}}$  is estimated to be -10.0 kcal/mole.

Thus, the microplate thermal shift assay offers many advantages for analyzing multi-ligand binding interactions in functional genomics classification

studies. For example, the same assay can simultaneously detect the binding of different kinds of ligands that bind at multiple binding sites on a target protein. Each ligand binding interaction identified aids the user in assigning a function to a protein. When the functions are summed up, one obtains a response curve that is characteristic of a particular class of proteins.

For example, if one considers the information obtained here for thrombin, and for the moment forget what is known about this protein, the heparin binding data might suggest an extracellular role for this protein since heparin and other sulfated oligosaccharide are important components of the extracellular matrix of the tissues of higher organisms. The catalytic binding site ligand, 3DP-4660, is a non-peptide mimic of a peptide that has an arginyl side chain at the P1 position, characteristic of substrates and inhibitors of trypsin-like serine proteases. Similarly, boroarginine transition state analogs, which have an arginine group in the P1 position for this synthetic peptide mimic, were found to be specific inhibitors for the serine proteases, thrombin, trypsin, and plasmin (Tapparelli *et al.*, *J. Biol. Chem.* 268:4734-4741 (1993)) with the observed specificity:  $K_d \sim 10$  nM (thrombin),  $K_d \sim 1,000$  nM (trypsin),  $K_d \sim 10,000$  nM (plasmin). Thus, the combined knowledge of heparin binding with the observed binding to boroarginine transition state analogs would quickly focus the assignment of this protein to an extracellular proteolytic function in the absence of any other information.

Further, the microplate thermal shift assay can be used, in a high throughput fashion, to detect cooperativity in ligand binding. Information about ligand binding cooperativity can be collected and analyzed very quickly, over a few hours, rather than over several months, as is required when conventional methods are used to classify protein function.

### ***Example 3***

#### ***Functional Probe Library Screen against Human Factor Xa***

A functional probe library is shown in Figure 5. A 96 well plate (Plate 1) contained 94 compounds (and two control wells) and included many compounds that are considered useful for providing information about the ligand binding preferences, and thus probable function, of proteins. For example, cofactors such as NAD and ATP are found in wells A4 and A5, respectively. This particular plate also contained a great many metal ion binding conditions to help probe a target protein for metal ion cofactors.

In order to validate the functional probe screen, two known proteins were incubated with the compounds of Plate 1 and were then assayed using the microplate thermal shift assay. For example, the activity spectrum obtained for Factor Xa (Enzyme Research Labs) is shown in Figure 6.

Factor Xa was purchased from Enzyme research Labs (South Bend, IN). Reactions were prepared in 96-well polycarbonate microtitre plate v-bottom wells. The final concentration of Factor Xa was 1.4  $\mu$ M (55 ng/mL) in 200 mM Tris-HCl, pH 8. The final concentration of 1,8-ANS was 100  $\mu$ M. The final concentration of each of the molecules tested for binding is shown in Figure 6. The contents were mixed by repeated uptake and discharge in a 100  $\mu$ L pipette tip. Finally, one drop of mineral oil (Sigma, St. Lois, MO) was added on top of each reaction well to reduce evaporation from samples at elevated temperatures.

The microplate reactions were heated simultaneously, in two degree increments, from 40 to 70° C, using a RoboCycler® Gradient Temperature Cycler (Stratagene, La Jolla, CA). After each heating step, prior to fluorescence scanning, the sample was cooled to 25°C. Fluorescence was measured using a CytoFluor II fluorescence microplate reader (PerSeptive Biosystems, Framingham, MA). 1,8-ANS was excited with light at a wavelength of 360 nm. The fluorescence emission was measured at 460 nm.

There were found to be six conditions that stabilized this enzyme with a  $\Delta T_m$  of greater than 1.0°C: (1) 0.5 M  $(\text{NH}_4)_2\text{SO}_4$ , (2) 0.5 M  $\text{MgSO}_4$ , (3) 0.5 M  $\text{Li}_2\text{SO}_4$ , (4) 0.5 M  $\text{KCl}$  (5) 0.1 M tri-polyphosphate, and (6) 0.1 M  $\text{CaCl}_2$ . The last two conditions are probably most significant, since tri-polyphosphate is a polyelectrolyte that mimics heparin and other sulfated oligosaccharides, and its binding to proteins suggests the presence of a heparin binding site, something that is well known for Factor Xa. Similarly,  $\text{Ca}^{2+}$  is known to bind to the Gla domain of Factor Xa, which is consistent with the stabilizing effect seen for 0.1 M  $\text{CaCl}_2$ .

Some of metal ions were found to have a strong destabilizing effect on Factor Xa. For example,  $[\text{Co}(\text{NH}_3)_6]\text{Cl}_3$ ,  $\text{BaCl}_2$ ,  $\text{CdCl}_2$ ,  $\text{YCl}_3$ , and  $\text{NiSO}_4$  were observed to destabilize Factor Xa by from 6 to 17°C. The reason for this destabilizing effect is unknown. It is possible that these metal ions preferentially bind to the unfolded form of Factor Xa. Some interference with the fluorescence probe is also possible.

#### Example 4

##### *Functional Probe Library Screen against Human D(II) FGFR1.*

The compounds in functional probe library Plate 1 was also employed to generate an activity spectrum for D(II) FGFR1. D(II) FGFR1 was cloned and expressed in *E. coli*. Recombinant D(II) FGFR1 was renatured from inclusion bodies essentially as described (Wetmore, D.R. *et al.*, Proc. Soc. Mtg., San Diego, CA (1994)), except that a hexa-histidine tag was included at the N-terminus to facilitate recovery by affinity chromatography on a  $\text{Ni}^{2+}$  chelate column (Janknecht, R. *et al.*, *Natl. Acad. Sci. USA* 88:8972-8976 (1991)). D(II) FGFR1 was further purified on a heparin-sepharose column (Kan, M. *et al.*, *Science* 259:1918-1921 (1993); Pantoliano, M.W. *et al.*, *Biochemistry* 33:10229-10248 (1994)). Purity was >95%, as judged by SDS-PAGE. The D(II) FGFR1 protein was concentrated to 12 mg/mL (~1 mM) and stored at 4°C.

Reactions were prepared in 96-well polycarbonate microtitre plate v-bottom wells. The final concentration of D(II) FGFR1 was 50  $\mu$ M in 200 mM Tris-HCl, pH 8 in each well of a 96-well polycarbonate microtitre plate. The final concentration of 1,8-ANS was 100  $\mu$ M. The final concentration of each of the molecules tested for binding is shown in Figure 7. The contents were mixed by repeated uptake and discharge in a 100  $\mu$ L pipette tip. Finally, one drop of mineral oil (Sigma, St. Lois, MO) was added on top of each reaction well to reduce evaporation from samples at elevated temperatures.

The microplate reactions were heated simultaneously, in two degree increments, from 25 to 60° C, using a RoboCycler® Gradient Temperature Cycler (Stratagene, La Jolla, CA). After each heating step, prior to fluorescence scanning, the sample was cooled to 25°C. Fluorescence was measured using a CytoFluor II fluorescence microplate reader (PerSeptive Biosystems, Framingham, MA). 1,8-ANS was excited with light at a wavelength of 360 nm. The fluorescence emission was measured at 460 nm.

The resultant activity spectrum is shown in Figure 7. A larger number of compounds were found to stabilize D(II) FGFR1. For example, all of the sugars, D(+)-glucose, D(+)-sucrose, xylitol, and sorbitol were all found to stabilize (and presumably bind) to D(II) FGFR1. This result may be consistent with the known heparin binding properties of this protein. Tri-polyphosphate, a known polyelectrolyte heparin mimic, yielded the largest shift: about 11°C. This result is consistent with the heparin binding properties of this protein (Pantoliano, M.W. *et al.*, *Biochemistry* 33:10229-10248 (1994)).

Thus, in a situation where a user did not know anything about this protein (as is typically the case when a new gene is cloned and the function of the encoded protein is unknown), the information obtained by screening just the compounds in Plate 1 would have provided a user some evidence that D(II)FGFR1 could be classified as a heparin-binding protein.

### Example 5

#### *Identification of Protein Targets containing DNA Binding Sites*

The *lac* repressor is normally tetrameric protein, a dimer of dimers. However, this protein has been shown to bind to DNA in its dimeric state. Lewis *et al.* solved the crystal structure of Lac repressor bound to its cognate DNA ligand (Lewis *et al.*, 1996, *Science* 271:1247-1254). A genetically altered dimer, one that is unable to form a tetramer, and a synthetic 21-mer oligonucleotide, the palindromic sequence of the native *lac* operator, were obtained from Dr. Mitch Lewis at the University of Pennsylvania. Binding of the synthetic *lac* operator to the mutant *lac* repressor was assayed using the microplate thermal shift assay.

The final concentration of *lac* repressor was 60  $\mu$ M in 200 mM Tris·HCl, pH 8. Reactions were prepared in 96-well polycarbonate microtiter plate V-bottom wells. The final concentration of 1,8-ANS was 100  $\mu$ M. The final concentration of each of the molecules tested for binding is shown in Figure 7. The contents were mixed by repeated uptake and discharge in a 100  $\mu$ L pipette tip. Finally, one drop of mineral oil (Sigma, St. Lois, MO) was added on top of each reaction well to reduce evaporation from samples at elevated temperatures.

The microplate reactions were heated simultaneously, in two degree increments, from 25°C to 75°C, using a ROBOCYCLER® Gradient Temperature Cycler (Stratagene, La Jolla, CA). After each heating step, prior to fluorescence scanning, the sample was cooled to 25°C. Fluorescence was measured using a CytoFluor II fluorescence microplate reader (PerSeptive Biosystems, Framingham, MA). ANS was excited with light at a wavelength of 360 nm. The fluorescence emission was measured at 460 nm.

In the presence of 80  $\mu$ M synthetic operator DNA, the  $T_m$  for the unfolding transition of *lac* repressor was shifted 5.6°C (Figure 8). The calculated  $K_d$  at  $T_m$  is 6  $\mu$ M. Using educated guesses for  $\Delta H_L$  (-10.0 kcal/mol), the calculated  $K_d$  at 25°C is 1.2  $\mu$ M and the calculated  $K_d$  at physiological temperature (37°C) is 3.4  $\mu$ M. The fluorescent probe, 1,8-ANS, did not bind to

DNA alone (*i.e.*, there was no fluorescence signal for the control reaction in which no *lac* repressor was included).

These results show that the microplate thermal shift assay can be used to assay DNA/protein interactions.

5

### ***Example 6***

#### ***Assays of ATP Binding***

10

15

Adenosine triphosphate (ATP) and ATP analogue binding can be assayed using the microplate thermal shift assay. Bovine muscle myosin (Sigma), bovine heart 3'-5' cAMP-dependent protein kinase (Sigma), and chicken muscle pyruvate kinase (Sigma) were each dissolved in Buffer A to generate stock solutions at a final concentration of 2 mg/mL. Magnesium chloride ( $\text{MgCl}_2$ ), adenosine triphosphate, adenosine triphosphate- $\gamma$ -S (ATP- $\gamma$ -S), aluminum trifluoride ( $\text{AlF}_3$ ), and sodium fluoride (NaF) were dissolved in Buffer A (50 mM HEPES, pH 7.5, 100 mM NaCl) to the stock concentrations used in each experiment. Dapoxyl™ 12800 solution was prepared by diluting a stock of 20 mM Dapoxyl 12800™ (5-(4"-dimethylaminophenyl)-2-(4'-phenyl)oxazole sulfonic acid, sodium salt, Molecular Probes, Inc.) in dimethyl sulfoxide to the appropriate concentration in Buffer A.

20

25

In the ATP and ATP- $\gamma$ -S reactions, each sample contained 12  $\mu\text{L}$  of protein stock solution (2 mgs/mL), 9.6  $\mu\text{L}$  of either ATP or ATP- $\gamma$ -S (50 mM), 4.8  $\mu\text{L}$  of  $\text{MgCl}_2$  (100 mM) and 21.6  $\mu\text{L}$  of a solution of 222  $\mu\text{M}$  of dapoxyl 12800 in Buffer A. In the ATP, aluminum trifluoride, and sodium fluoride reactions, each sample contained 12  $\mu\text{L}$  of protein stock solution (2 mgs/mL), 9.6  $\mu\text{L}$  of ATP (50 mM), 9.6  $\mu\text{L}$  of aluminum trifluoride (50 mM) + sodium fluoride (50 mM), 4.8  $\mu\text{L}$  of 100 mM  $\text{MgCl}_2$ , and 12  $\mu\text{L}$  of a solution of 400  $\mu\text{M}$  Dapoxyl 12800 in Buffer A.

For the thermal shift assay, four 10  $\mu\text{L}$  aliquots of each assay mixture were dispensed into four wells located in different quadrants of an MJ Research 384-



well thermocycler plate. 10  $\mu$ L of mineral oil was then added to each of the four wells to prevent evaporation. Each data point shown was collected by heating the plate at the temperatures shown for three minutes. For example, the plate was heated to a given temperature, and then allowed to cool to 25°C for one minute, followed by UV illumination and collection of the data. Then the plate was heated to the next higher temperature, and so forth. UV illumination was performed using a long wavelength illumination at 200–420 nm, having a peak at 365 nm. Fluorescence was imaged using a CCD camera having a bandpass filter centered at 550 nm.

Figure 9 shows the results of a microplate thermal shift assay of ATP to bovine muscle myosin. The data is plotted as fluorescence intensity as a function of temperature. The  $T_m$  of the control thermal unfolding curve (no ATP) was 49.3°C (microplate well K2). The  $T_m$  of the thermal unfolding curve for bovine muscle myosin bound to ATP ((+) ATP) was 51.4°C (microplate well K16). Thus the  $\Delta T_m$  for ATP binding was 2.1°C. The  $K_d$  was 440  $\mu$ M.

Figure 10 shows the result of a microplate thermal shift assay of ATP- $\gamma$ -S to and 3', 5'-cAMP-dependent protein kinase. The data is plotted as fluorescence intensity as a function of temperature. The  $T_m$  of the control thermal unfolding curve (no ATP- $\gamma$ -S) was 46.2°C (microplate well E14). The  $T_m$  of the thermal unfolding curve for 3', 5'-cAMP-dependent protein kinase bound to ATP- $\gamma$ -S ((+) ATP- $\gamma$ -S) was 51.8°C (microplate well M15). Thus the  $\Delta T_m$  for ATP- $\gamma$ -S binding was 5.6°C. The  $K_d$  was 200  $\mu$ M. The results, including the results for pyruvate kinase, are summarized in Table 4.

**Table 4. Summary of results for enzymes that bind to ATP. The value in parentheses is standard deviation.**

Protein	Reference $T_m$	ATP- $\gamma$ -S (10 mM)	ATP (10 mM)	ATP + $AlF_3$ (10 mM)
		$\Delta T_m$	$\Delta T_m$	$\Delta T_m$
Myosin	49.4	0.0 ( $\pm 0.2$ )	2.2 ( $\pm 0.4$ )	2.8 ( $\pm 0.4$ )
3' - 5' cAMP Protein kinase	44.7	5.6 ( $\pm 1.7$ )	7.5 ( $\pm 0.7$ )	8.2 ( $\pm 1.4$ )
Pyruvate kinase	54.5	0.8 ( $\pm 0.11$ )	-0.44 ( $\pm 0.1$ )	-0.27 ( $\pm 0.2$ )

### *Example 7*

#### *Assay of Folic Acid Binding*

Folic acid binding can be assayed using the microplate thermal shift assay. Bovine liver dihydrofolate reductase (DHFR, Sigma), chicken liver dihydrofolate reductase (DHFR, Sigma), pigeon liver arylamine acetyltransferase (ArAcT, Sigma), and porcine liver formimino glutamic acid transferase (FGT, Sigma) were each dissolved in Buffer A (50 mM HEPES, pH 7.5, 100 mM NaCl) to generate stock solutions at a final concentration of 2 mg/mL. Solutions of dihydrofolic acid ( $FAH_2$ ), methotrexate, nicotinamide adenine dinucleotide phosphate (NADP), were prepared by dissolving solid material into Buffer A immediately prior to use. Dapoxyl™ 12800 solution was prepared by diluting a stock of 20 mM Dapoxyl™ 12800 in dimethyl sulfoxide to the appropriate concentration in Buffer A.

Each assay sample contained 12  $\mu$ L of protein stock solution (2 mg/mL), 4.8  $\mu$ L of either dihydrofolic acid ( $FAH_2$ ) or methotrexate stock solution (1 mM), and 31.2  $\mu$ L of a solution of 154  $\mu$ M Dapoxyl™ 12800 in Buffer A. Each sample contained 12  $\mu$ L of protein stock solution (2 mgs/mL), 4.8  $\mu$ L of NADP stock

solution (50 mM), and 31.2  $\mu$ L of a solution of 154  $\mu$ M Dapoxyl™ 12800 in Buffer A.

For the thermal shift assay, four 10  $\mu$ L aliquots of each assay mixture were dispensed into four wells located in different quadrants of an MJ Research 384-well thermocycler plate. 10  $\mu$ L of mineral oil was then added to each of the four wells to prevent evaporation. Each data point shown was collected by heating the plate at the temperature shown for three minutes, followed by incubation at 25°C for one minute, followed by UV illumination and collection of the data. The results are shown in Table 5.

**Table 5. Results for proteins that bind methotrexate, FAH<sub>2</sub> and NADP. The value in parentheses is standard deviation.**

Protein	Reference $T_m$	Methotrexate (100 $\mu$ M) $\Delta T_m$	FAH <sub>2</sub> (100 $\mu$ M) $\Delta T_m$	NADP (5 mM) $\Delta T_m$
DHFR	52.47	7.0 ( $\pm$ 0.1)	-0.64 ( $\pm$ 0.2)	3.2 ( $\pm$ 0.13)
DHFR	56.6	8.6 ( $\pm$ 0.2)	2.5 ( $\pm$ 0.2)	3.8 ( $\pm$ 0.4)
Arylamine Acetyl- transferase	49.8	1.0 ( $\pm$ 0.4)	-1.8 ( $\pm$ 0.5)	2.8 ( $\pm$ 0.4)
Formimino L-Glutamic acid Transferase	47.2	0.9 ( $\pm$ 0.5)	3.62 ( $\pm$ 0.4)	0.0 ( $\pm$ 0.2)

### *Example 8*

#### *Assay of Methotrexate/NADP(H) Binding*

The ability to measure temperature shifts for the binding of methotrexate and NADPH, both separately and simultaneously, is another example of the utility of the present invention in measuring multi-ligand binding interactions. In this case, the binding sites of the two ligands are proximal, and there is positive cooperativity in the binding of the two ligands, as shown by the fact that thermal shift for both ligands binding simultaneously is 2-4 degrees more than the total of shifts for each ligand binding separately (Table 6).

Methotrexate (MTX) and NADPH binding can be assayed using the microplate thermal shift assay. Bovine liver dihydrofolate reductase (DHFR, Sigma) and chicken liver dihydrofolate reductase (DHFR, Sigma) were each dissolved in Buffer A (50 mM HEPES, pH 7.5, 100 mM NaCl) to generate stock solutions at a final concentration of 2 mg/mL. All stock solutions of ligands were prepared by dissolving solid material in Buffer A immediately prior to use. Stock solutions of nicotinamide adenine dinucleotide phosphate-reduced form (NADPH, 100 mM), NADP (100 mM), and methotrexate (1 mM) were diluted further in Buffer A to twice to the final assay concentration (2x stocks): methotrexate (200  $\mu$ M), NADP (20 mM), NADPH (20 mM), methotrexate + NADP (200  $\mu$ M + 20 mM), methotrexate + NADPH (200  $\mu$ M + 20 mM). Dapoxyl™ 12800 solutions was prepared by diluting a stock of 20 mM Dapoxyl™ 12800 in dimethyl sulfoxide to the appropriate concentration in Buffer A. 5  $\mu$ L of each protein stock solution was added to 25  $\mu$ L of 2x ligand stock solution mixed with 20  $\mu$ L of a solution of 250  $\mu$ M Dapoxyl™ 12800 in Buffer A.

The final ligand concentrations were 10 mM NADP; 10 mM NADPH; and 100  $\mu$ M MTX.

For the thermal shift assay, four 10  $\mu$ L aliquots of each assay mixture were dispensed into four wells located in different quadrants of an MJ Research 384-well thermocycler plate. 10  $\mu$ L of mineral oil was then added to each of the four

wells to prevent evaporation. Each data point shown was collected by heating the plate at the temperature shown for three minutes, followed by incubation at 25°C for one minute, followed by UV illumination and collection of the data.

Figure 11 shows the result of a microplate thermal shift assay of methotrexate to dihydrofolate reductase. The data is plotted as fluorescence intensity as a function of temperature. The  $T_m$  of the control thermal unfolding curve (no MTX) was 47.2°C (microplate well M1). The  $T_m$  of the thermal unfolding curve for DHFR bound to methotrexate ((+) MTX) was 56.4°C (microplate well G6). Thus the  $\Delta T_m$  for methotrexate binding was 9.2°C. The  $K_d$  was 24 nM.

Figure 12 shows the result of a microplate thermal shift assay of NADPH to dihydrofolate reductase. The data is plotted as fluorescence intensity as a function of temperature. The  $T_m$  of the control thermal unfolding curve (no NADPH) was 50.8°C (microplate well G8). The  $T_m$  of the thermal unfolding curve for DHFR bound to NADPH ((+) NADPH) was 53.8°C (microplate well B20). Thus the  $\Delta T_m$  for NADPH binding was 3.0°C. The  $K_d$  was 0.7  $\mu$ M.

**Table 6.  $\Delta T_m$ 's of ligand complexed with DHFR. The value in parentheses is standard deviation.**

Protein	NADP	MTX	Sum <sup>a</sup>	NADP + MTX <sup>b</sup>	NADP H	MTX	Sum <sup>a</sup>	NADPH + MTX <sup>b</sup>
DHFR (chicken)	7.5 ( $\pm 0.38$ )	10.1 ( $\pm 0.32$ )	17.6	20.9 ( $\pm 0.4$ )	11.9 ( $\pm 1.3$ )	10.1 ( $\pm 0.32$ )	22	23.8 ( $\pm 0.4$ )
DHFR (cow)	6.3 ( $\pm 0.1$ )	7.7 ( $\pm 0.3$ )	14	18.1 ( $\pm 0.4$ )	9.7 ( $\pm 0.2$ )	7.7 ( $\pm 0.3$ )	17.4	24.6 ( $\pm 0.6$ )

<sup>a</sup> The value shown is the sum of the individual  $\Delta T_m$ 's observed from the protein incubated separately with each ligand.

<sup>b</sup> The value shown is the  $\Delta T_m$  observed when the protein was incubated simultaneously with both ligands.

### *Example 9*

Dihydrofolic acid is a substrate of dihydrofolate reductase (DHFR). Methotrexate is a folic acid analog that binds to DHFR. As evidence that the method of the present invention is reliable, it was shown that the method can be used to detect binding of dihydrofolic acid to DHFR. Bovine liver DHFR was combined with 80 compounds to screen for the function of the protein, and binding to methotrexate, but not to a number of other compounds, was detected.

Each well of microsource compound plate #198104 contained one of 80 different compounds at a concentration 10 mM in dimethyl sulfoxide. Each compound solution was diluted in Buffer A (50 mM HEPES, pH 7.5, 100 mM NaCl) to a final concentration of 200  $\mu$ M in separate wells in a 384-well polystyrene plate. 5  $\mu$ L of the solution contained in each well was transferred to an MJ research polypropylene plate containing 5  $\mu$ L of bovine liver DHFR (at a concentration of 0.5 mg/mL and Dapoxyl™ 12800 dye at a concentration of 200  $\mu$ M, yielding final concentrations of 100  $\mu$ M ligand, 0.25 mg/mL DHFR, and 100  $\mu$ M dapoxyl in the 10 L volume of each well.

10  $\mu$ L of mineral oil was added to each well to prevent evaporation. Thermal unfolding profiles were then measured for each well from 25 to 70°C, by collecting data points at each temperature, separated by one-degree increments. Each data point was collected by heating the plate at the temperature shown for 3 minutes, followed by incubation at 25°C for one minute, followed by long-wavelength UV illumination and collection of the data using a CCD camera.

The data were collected as four replicates of 80 compounds in the quadrants of a 384-well plate. The four quadrants consist of: wells A2 through H11 (first quadrant), wells A14 through H23 (second quadrant), wells I2 through P11 (third quadrant), and wells I14 through I23 (fourth quadrant). Columns 1, 12, 13, and 24 consist of reference wells containing only DHFR and dimethyl sulfoxide.

Wells F2, F14, N2, and N14 contained methotrexate. Binding was revealed by fitting software as a red well. Methotrexate shifted the  $T_m$  by  $5.13 \pm 0.19$  degrees (average of 4 quadrants), and the other compounds on the plate had little or no effect (shown as near-white wells). These results which indicated that DHFR binds methotrexate.

All publications and patents mentioned hereinabove are hereby incorporated in their entireties by reference.

While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be appreciated by one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention and appended claims.

*What Is Claimed Is:*

1. A method for functionally classifying a protein, said method comprising:

5 (a) screening one or more of a multiplicity of different molecules for their ability to modify the stability of said protein, wherein modification of the stability of the protein indicates that the molecule binds to the protein;

(b) generating an activity spectrum for said protein from said screening of step (a), wherein said activity spectrum reflects a subset of molecules, from said multiplicity of different molecules, that modify the stability of said protein and  
10 therefore are ligands that bind to the protein;

(c) comparing said activity spectrum for said protein to one or more functional reference spectrum lists; and

(d) classifying said protein according to the set of molecules in said multiplicity of different molecules that modify the stability of said protein.

15 2. The method of claim 1, wherein said screening step (a) comprises:

(a1) contacting said protein with one or more of said multiplicity of different molecules in each of a multiplicity of containers;

(a2) treating said protein in each of said multiplicity of containers to cause said protein to unfold;

20 (a3) measuring in each of said containers a physical change associated with the unfolding of said target molecule;

(a4) generating an unfolding curve for said target molecule for each of said containers; and

25 (a5) comparing each of said unfolding curves in step (d) to (1) each of said other unfolding curves and to (2) the unfolding curve obtained for said protein in the absence of any of said multiplicity of different molecules; and



(a6) determining whether any of said multiplicity of different molecules modifies the stability of said protein, wherein a modification in stability is indicated by a change in said unfolding curve.

3. A method for functionally classifying a protein, said method comprising:

(a) screening one or more of a multiplicity of different molecules known to bind to a particular class of proteins for their ability to modify the stability of said protein, wherein modification of the stability of the protein indicates that the molecule binds to the protein;

(b) generating an activity spectrum for said protein from said screening of step (a), wherein said activity spectrum reflects a subset of molecules, from said multiplicity of different molecules, that modify the stability of said protein and therefore are ligands that bind to the protein; and

(c) classifying said protein as a member of said class of proteins if said one or more of said multiplicity of different molecules modify the stability of said protein.

4. The method of claim 3, wherein said screening step (a) comprises:

(a1) contacting said protein with one or more of said multiplicity of different molecules in each of a multiplicity of containers;

(a2) treating said protein in each of said multiplicity of containers to cause said protein to unfold;

(a3) measuring in each of said containers a physical change associated with the unfolding of said target molecule;

(a4) generating an unfolding curve for said target molecule for each of said containers; and

(a5) comparing each of said unfolding curves in step (d) to (1) each of said other unfolding curves and to (2) the unfolding curve obtained for said protein in the absence of any of said multiplicity of different molecules; and

(a6) determining whether any of said multiplicity of different molecules modifies the stability of said protein, wherein a modification in stability is indicated by a change in said unfolding curve.

5 5. A method for functionally classifying a protein, said method comprising:

classifying said protein according to the set of molecules in a multiplicity of different molecules that modify the stability of said protein.

6. A method for functionally classifying a protein that is capable of unfolding due to a thermal change, said method comprising:

10 (a) screening one or more of a multiplicity of different molecules for their ability to shift the thermal unfolding curve of said protein, wherein a shift in the thermal unfolding curve of the protein indicates that the molecule binds to the protein;

15 (b) generating an activity spectrum for said protein from said screening of step (a), wherein said activity spectrum reflects a subset of molecules, from said multiplicity of different molecules, that shift the thermal unfolding curve of said protein and therefore are ligands that bind to the protein;

(c) comparing said activity spectrum for said protein to one or more functional reference spectrum lists; and

20 (d) classifying said protein according to the set of molecules in said multiplicity of different molecules that shift the thermal unfolding curve of said protein.

7. The method of claim 6, wherein said screening step (a) comprises:

25 (a1) contacting said protein with one or more of said multiplicity of different molecules in each of a multiplicity of containers;

(a2) heating said multiplicity of containers from step (a1);

(a3) measuring in each of said containers a physical change associated with the thermal unfolding of said target molecule resulting from said heating;

(a4) generating a thermal unfolding curve for said target molecule as a function of temperature for each of said containers; and

5 (a5) comparing each of said unfolding curves in step (a4) to (1) each of said other thermal unfolding curves and to (2) the thermal unfolding curve obtained for said protein in the absence of any of said multiplicity of different molecules; and

10 (a6) determining whether any of said multiplicity of different molecules shift the thermal unfolding curve of said protein.

8. The method of claim 7, wherein said comparing step (a5) comprises ranking said molecules in said multiplicity of different molecules for said protein according to the ability of each of said multiplicity of different molecules to shift the thermal unfolding curve of said protein.

15 9. The method of claim 7, wherein in said heating step (a2), said multiplicity of containers is heated simultaneously.

10. The method of claim 7, wherein said step (a4) further comprises determining a midpoint temperature ( $T_m$ ) from the thermal unfolding curve; and wherein said step (a5) further comprises comparing the  $T_m$  of each of said  
20 unfolding curves in step (a4) to (1) the  $T_m$  of each of said other thermal unfolding curves and to (2) the  $T_m$  of the thermal unfolding curve obtained for said target protein in the absence of any of said different molecules.

11. The method of claim 7, wherein said step (a3) comprises measuring the absorbance of light by said contents of each of said containers.

12. The method of claim 7, wherein said step (a1) comprises contacting said protein with a fluorescence probe molecule present in each of said multiplicity of containers and wherein said step (a3) comprises

5 (i) exciting said fluorescence probe molecule, in each of said multiplicity of containers, with light; and

(ii) measuring the fluorescence from each of said multiplicity of containers.

10 13. The method of claim 12, wherein said step (a3)(ii) further comprises measuring the fluorescence from each of said multiplicity of containers one container at a time.

14. The method of claim 12, wherein said step (a3)(ii) further comprises measuring the fluorescence from a subset of said multiplicity of containers simultaneously.

15 15. The method of claim 12, wherein said step (a3)(ii) further comprises measuring the fluorescence from each of said multiplicity of containers simultaneously.

16. The method of claim 7, wherein said step (a3) comprises  
(i) exciting tryptophan residues in said protein, in each of said multiplicity of containers, with light; and

20 (ii) measuring the fluorescence from each of said multiplicity of containers.

17. The method of claim 7, wherein said multiplicity of containers in step (a1) comprises a multiplicity of wells in a microplate.

18. A method for functionally classifying a protein that is capable of unfolding due to a thermal change, said method comprising:

5 (a) screening one or more of a multiplicity of different molecules known to bind to a particular class of proteins for their ability to shift the thermal unfolding curve of said protein, wherein a shift in the thermal unfolding curve of the protein indicates that the molecule binds to the protein;

(b) generating an activity spectrum for said protein from said screening of step (a), wherein said activity spectrum reflects a subset of molecules, from said multiplicity of different molecules, that shift the thermal unfolding curve of said protein and therefore are ligands that bind to the protein; and

10 (c) classifying said protein as a member of said class of proteins if said one or more of said multiplicity of different molecules shift the thermal unfolding curve of said protein.

19. The method of claim 18, wherein said screening step (a) comprises:

15 (a1) contacting said protein with one or more of said multiplicity of different molecules in each of a multiplicity of containers;

(a2) heating said multiplicity of containers from step (a1);

(a3) measuring in each of said containers a physical change associated with the thermal unfolding of said target molecule resulting from said heating;

20 (a4) generating a thermal unfolding curve for said target molecule as a function of temperature for each of said containers; and

(a5) comparing each of said unfolding curves in step (a4) to (1) each of said other thermal unfolding curves and to (2) the thermal unfolding curve obtained for said protein in the absence of any of said multiplicity of different molecules; and

25 (a6) determining whether any of said multiplicity of different molecules shift the thermal unfolding curve of said protein.

20. The method of claim 19, wherein said comparing step (a5) comprises ranking said molecules in said multiplicity of different molecules for said protein according to the ability of each of said multiplicity of different molecules to shift the thermal unfolding curve of said protein.

5 21. The method of claim 19, wherein in said heating step (a2), said multiplicity of containers is heated simultaneously.

22. The method of claim 19, wherein said step (a4) further comprises determining a midpoint temperature ( $T_m$ ) from the thermal unfolding curve; and wherein said step (a5) further comprises comparing the  $T_m$  of each of said  
10 unfolding curves in step (a4) to (1) the  $T_m$  of each of said other thermal unfolding curves and to (2) the  $T_m$  of the thermal unfolding curve obtained for said target protein in the absence of any of said different molecules.

23. The method of claim 19, wherein said step (a3) comprises measuring the absorbance of light by said contents of each of said containers.

15 24. The method of claim 19, wherein said step (a1) comprises contacting said protein with a fluorescence probe molecule present in each of said multiplicity of containers and wherein said step (a3) comprises

(i) exciting said fluorescence probe molecule, in each of said multiplicity of containers, with light; and

20 (ii) measuring the fluorescence from each of said multiplicity of containers.

25. The method of claim 24, wherein said step (a3)(ii) further comprises measuring the fluorescence from each of said multiplicity of containers one container at a time.

26. The method of claim 24, wherein said step (a3)(ii) further comprises measuring the fluorescence from a subset of said multiplicity of containers simultaneously.

5 27. The method of claim 24, wherein said step (a3)(ii) further comprises measuring the fluorescence from each of said multiplicity of containers simultaneously.

28. The method of claim 19, wherein said step (a3) comprises  
(i) exciting tryptophan residues in said protein, in each of said multiplicity of containers, with light; and  
10 (ii) measuring the fluorescence from each of said multiplicity of containers.

29. The method of claim 18, wherein said multiplicity of containers in step (a1) comprises a multiplicity of wells in a microplate.

15 30. A method for functionally classifying a protein capable of unfolding due to a thermal change, said method comprising:  
classifying said protein according to the set of molecules in a multiplicity of different molecules that shift the thermal unfolding curve of said protein.

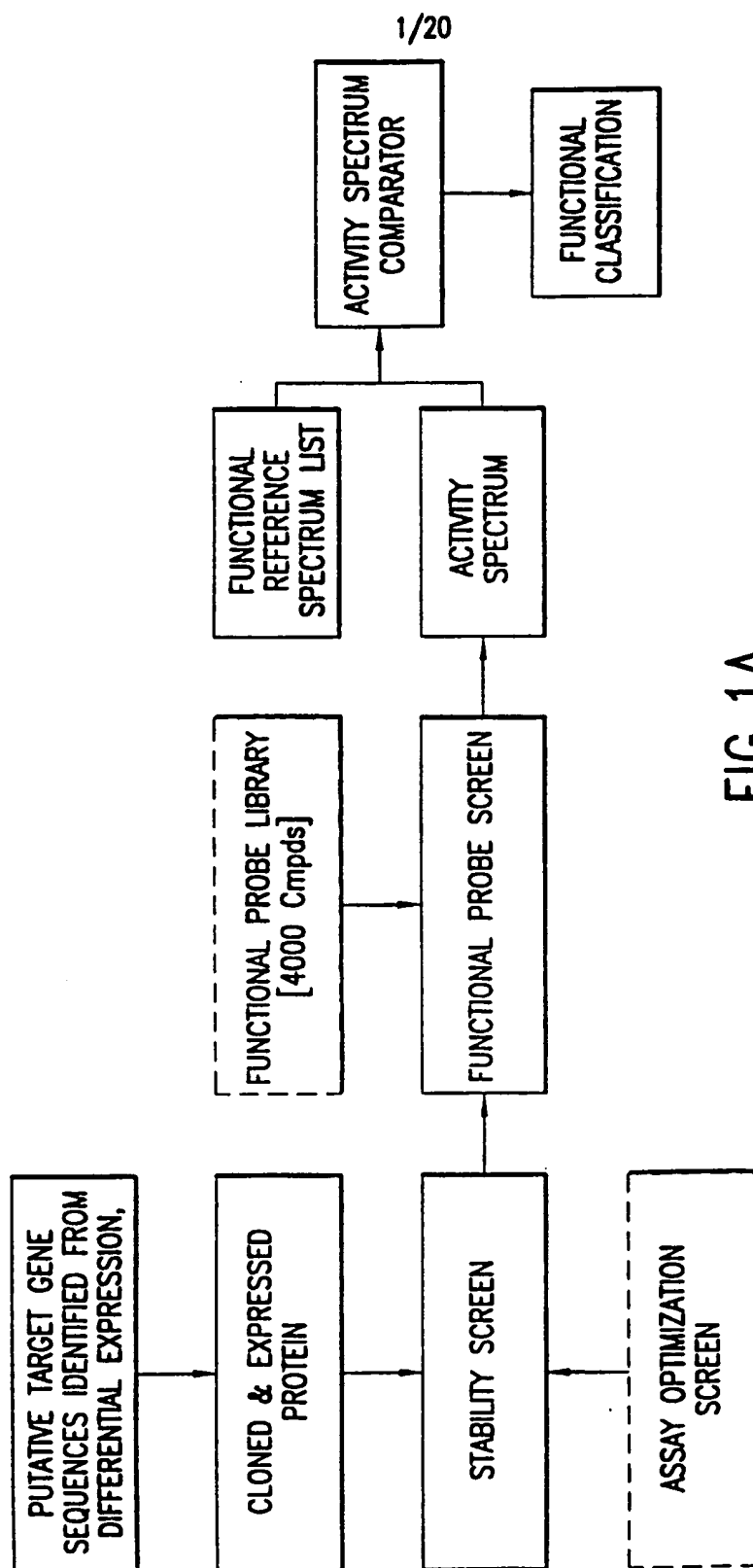


FIG. 1A



2/20

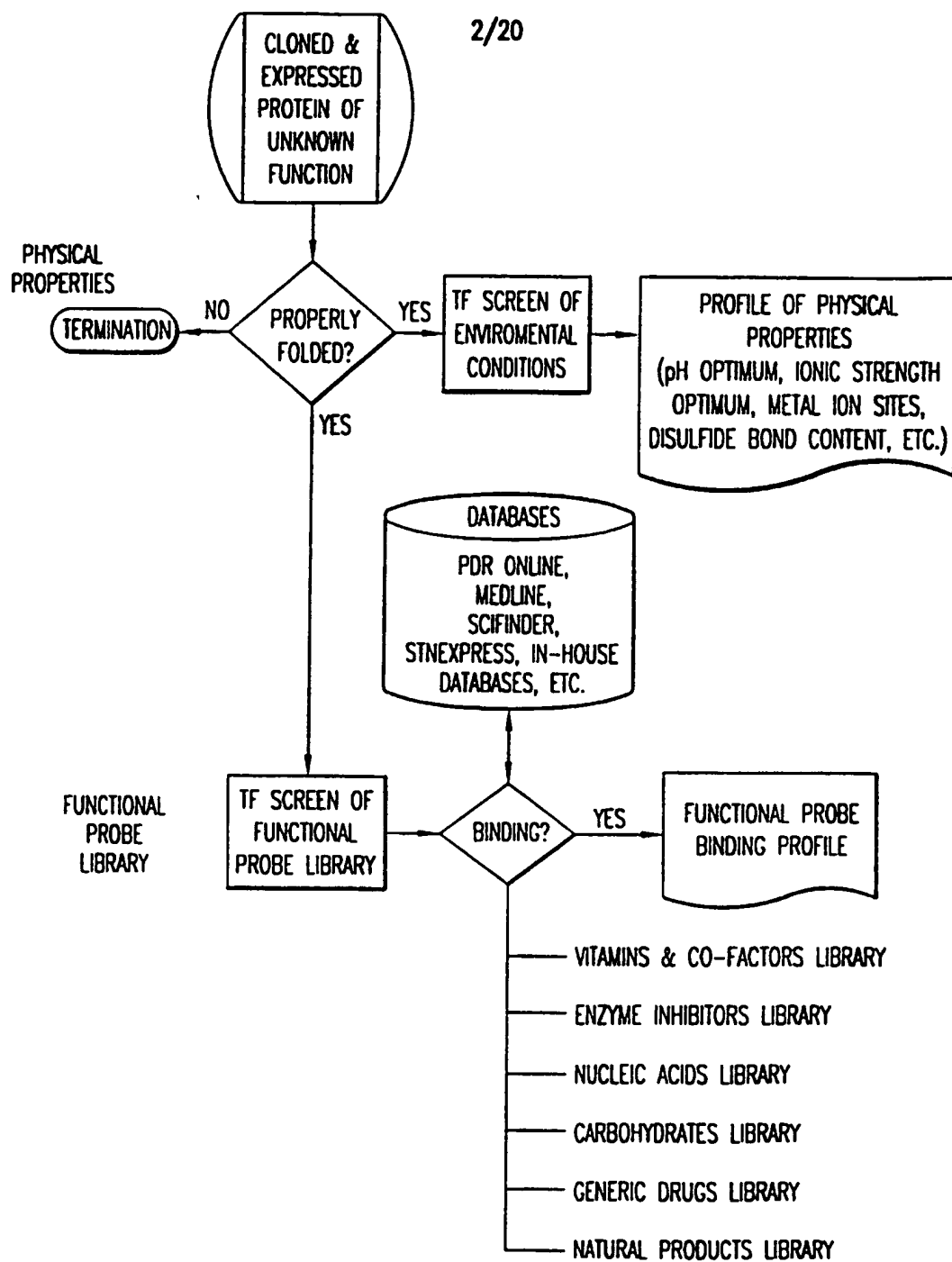


FIG. 1B

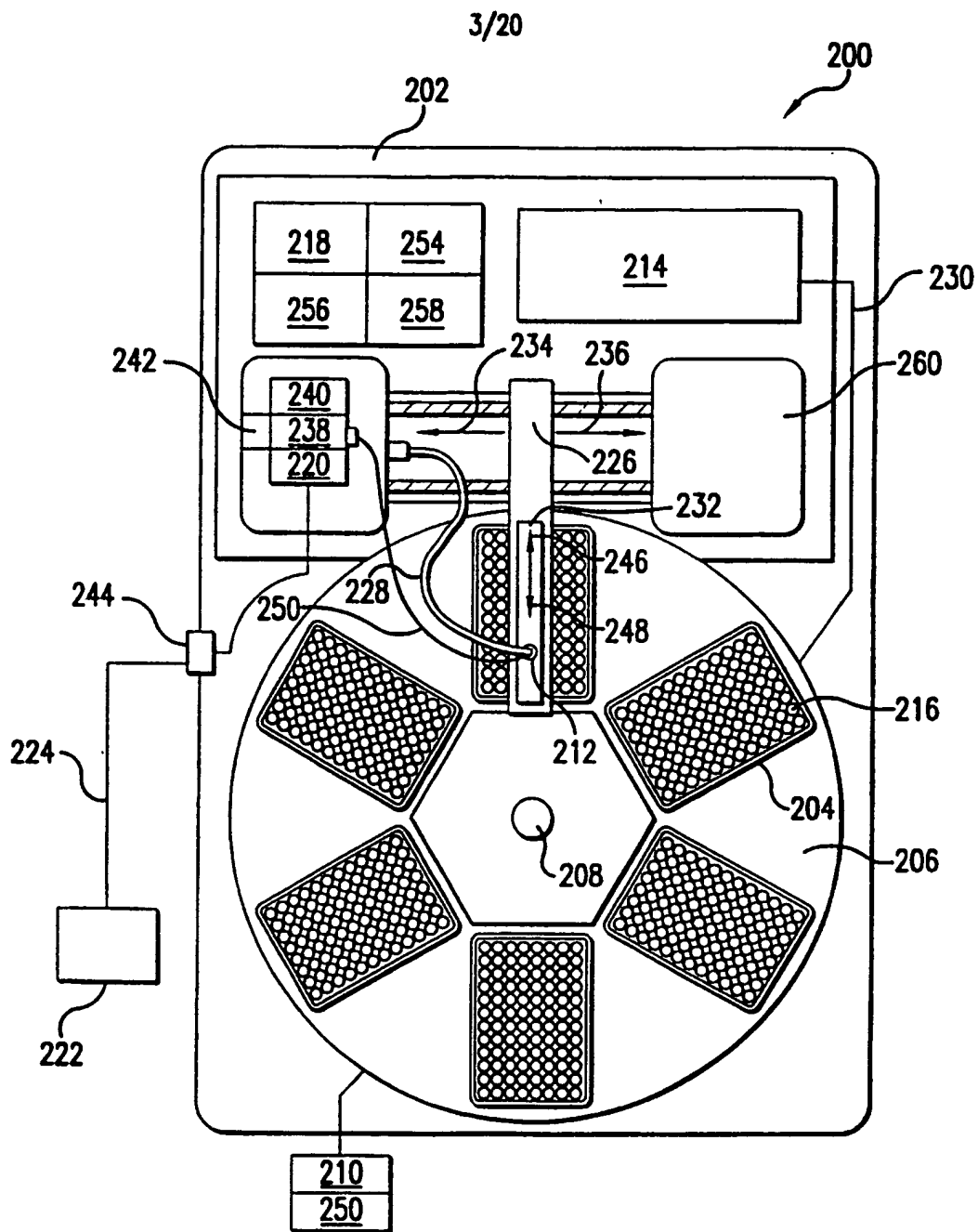


FIG.2

4/20

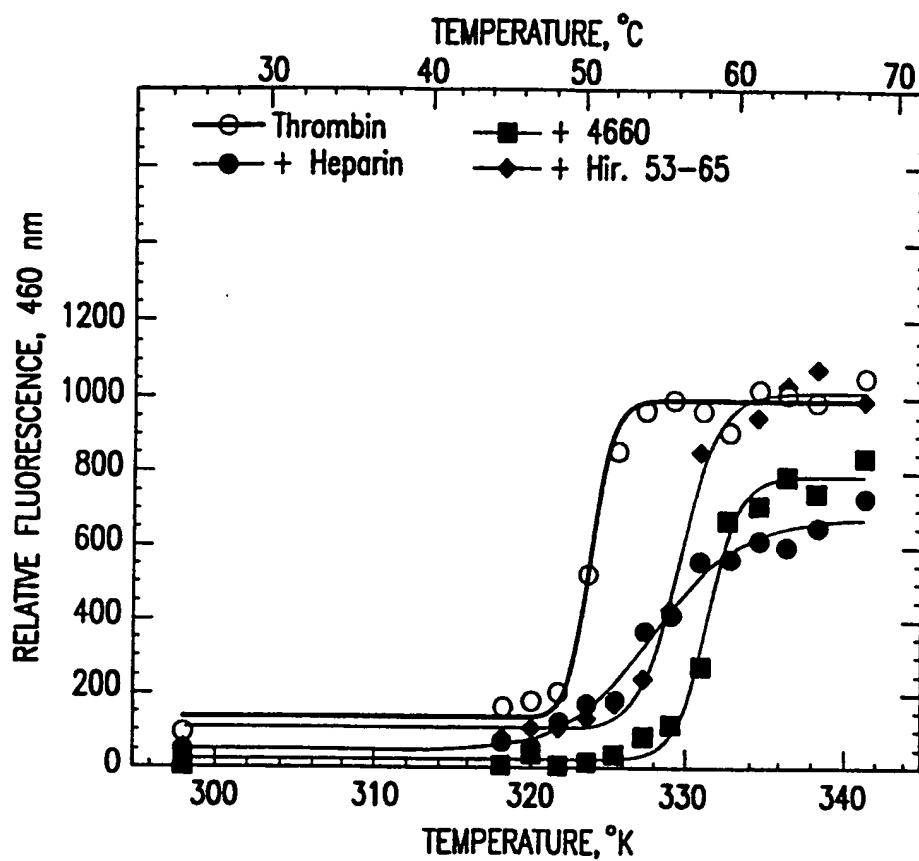


FIG.3

5/20

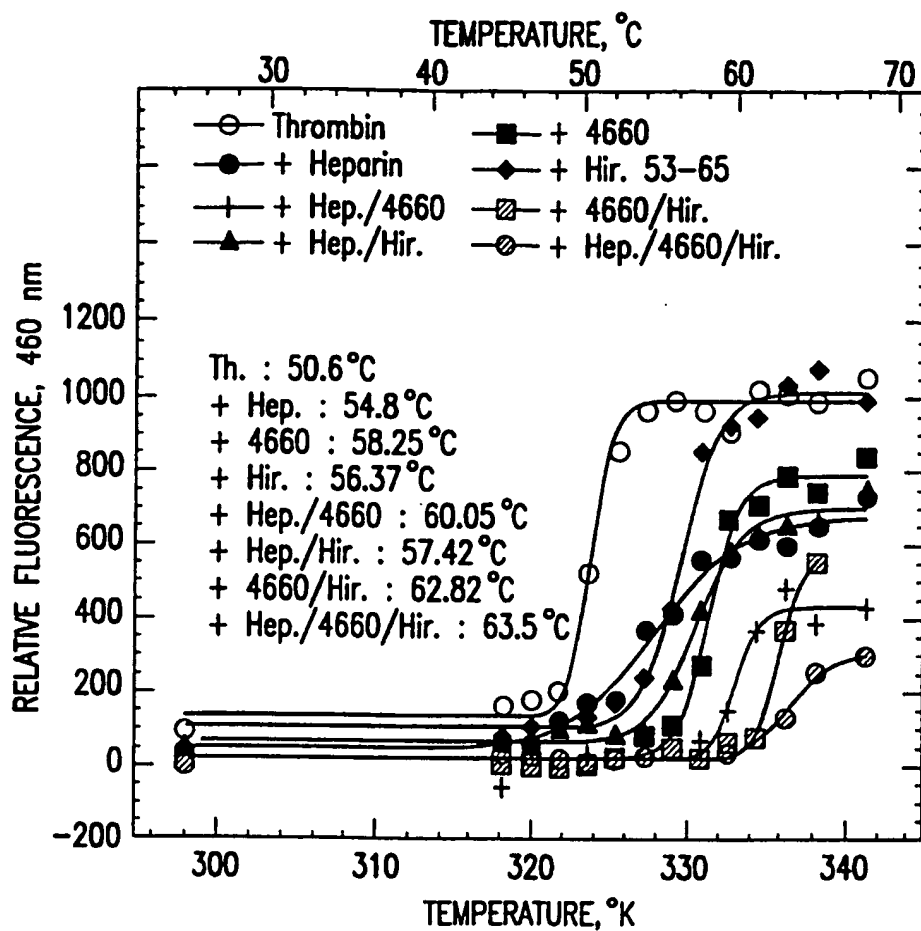


FIG.4

6/20

CONTINUED  
ON  
FIG.5B

	1	2	3	4	5	6
A	1. 0.1 M Na IODIDE?	2. 0.01 M L-CYSTEINE	3. 0.01 M EDTA	4. 0.01 M BETA-NICOTINAMIDE ADENOSINE DINUCLEOTIDE	5. 0.01 M ADENOSINE-5'- TRIPHOSPHATE DISODIUM SALT	6. 3% D(+)-GLUCOSE MONOHYDRATE
B	13. 3% 1,6-DIAMINOHEXANE	14. 3% 1,8-DIAMINOCTANE	15 0.1 M GLYCINE	16. 0.01 M Glycyl-glycyl-glycin	17. 0.01 M HEXAMINECOBALT TRICHLORIDE	18. 0.01 M TAURINE
C	25. 0.01 M Ba CHLORIDE	26. 0.01 M Cd CHLORIDE	27. 0.01 M CA CHLORIDE DIHYDRATE	28. 0.01 M COBALTOUS CHLORIDE	29. 0.01 M CUPRIC CHLORIDE DIHYDRATE	30. 0.01 M Mg CHLORIDE HEXAHYDRATE
D	37. 3% ETHELENE GLYCOL	38. 3% GLYCEROL	39. 3% 1,6 HEXANEDIOL	40. 3% ISOPROPANOL	41. 3% METHANOL	42. 3% MPD

CONTINUED ON FIG.5C

FIG. 5A

7/20

7	8	9	10	11	12
7. 3% D(+)-SUCROSE	8. 3% ZYLITOL	9. 0.01 M SPERMIDINE	10. 0.01 M SPERMINE TETRA-HCl	11. 3% 6-AMINOCAPROIC ACID	12. 3% 1,5-DIAMINOPENTAN di-HCl
19. 0.01 M BETAINE MONOHYDRATE	20. 0.5% POLYVINYLPYRROLIDONE	21. 0.3 M NON-DETERGENT SULFO-BETAINE 195	22. 0.2 M NON-DETERGENT SULFO-BETAINE 201	23. 3% DIMETHYL SULFOXIDE	24. 0.01 M PHENOL
31. 0.01 M Mn (11) CHLORIDE TETRAHYDRATE	32. 0.01 M STRONTIUM CHLORIDE 1	33. 0.01 M YTTRIUM CHLORIDE HE;	34. 0.01 M ZINC CHLORIDE?	35. 3% DIOXANE	36 3% ETHANOL
43. 5% PEG 400	44. 0.01 M TRIMETHYLAMINE HCl	45. 0.1 M GUANIDINE HCl	46. 0.01 M UREA	47. 1.5% 1,2,3-HEPTANETRIOL	48. 2% BENZAMIDINE HCl

CONTINUED  
FROM  
FIG. 5A

CONTINUED ON FIG. 5D

FIG. 5B

1	2	3	4	5	6

CONTINUED FROM FIG.5A

E	49. 0.1 M EGTA	50. 0.1 M NH <sub>4</sub> SULFATE	51. 0.5 M NH <sub>4</sub> SULFATE	52. 0.1 M NH <sub>4</sub> PHOSPHATE	53. 0.1 M NH <sub>4</sub> ACETATE	54. 0.1 M MG SULFATE
F	61. 0.5 M Li SULFATE	62. 0.1 M Ni SULFATE	63. 0.1 M K TARTRATE	64. 0.1 M K CHLORIDE	65. 0.5 M K CHLORIDE	66. 0.1 M K PHOSPHATE
G	73. 0.1 M Na PHOSPHATE	74. 0.5 M Na PHOSPHATE	75. 0.1 M Na CACODYLATE	76. 0.01 M di-Na PYROPHOSPHATE	77. 0.1 M Na TRI-METAPHOSPHATE	78. 0.1 M TRI-POLYPHOSPHATE
H	85. 0.1 M ARGININE	86. 0.01 M GLUTATHIONE	87. 0.1 M 3-NITROBENZENESULF ACID	88. 0.01 M DHDT	89. 0.01 M CYCLOHEXYLAMINE	90 0.01 M PIPERIDINE

CONTINUED  
ON  
FIG.5D

FIG.5C

7	8	9	10	11	12

CONTINUED FROM FIG.5B

55. 0.5 M Mg SULFATE	56. 0.1 M Mg CHLORIDE	57. 0.1 M Mg ACETATE	58. 0.1 M Li CHLORIDE	59. 0.5 M Li CHLORIDE	60. 0.1 M Li SULFATE
67. 0.2 M K SULFATE	68. 0.05 M K SULFATE	69. 0.1 M Na FORMATE	70. 0.1 M Na ACETATE	71. 0.1 M Na CHLORIDE	72. 0.1 M Na CITRATE
79. 0.1 M OXALATE	80. 0.1 M MALONATE	81. 0.1 M SUCCINATE	82. 0.1 M Zn ACETATE	83. 0.1 M SORBITOL	84. 0.1 M Ca CHLORIDE
91 0.01 M MORPHOLINE	92 0.005 M BENZOIC ACID	93 0.01 M BENZENESULFONIC ACID	94. 0.01 M SULFOBENZOIC ACID	95. CONTROL?	96. CONTROL?

CONTINUED  
FROM  
FIG.5A

E	F	G	H
---	---	---	---

FIG.5D



10/20

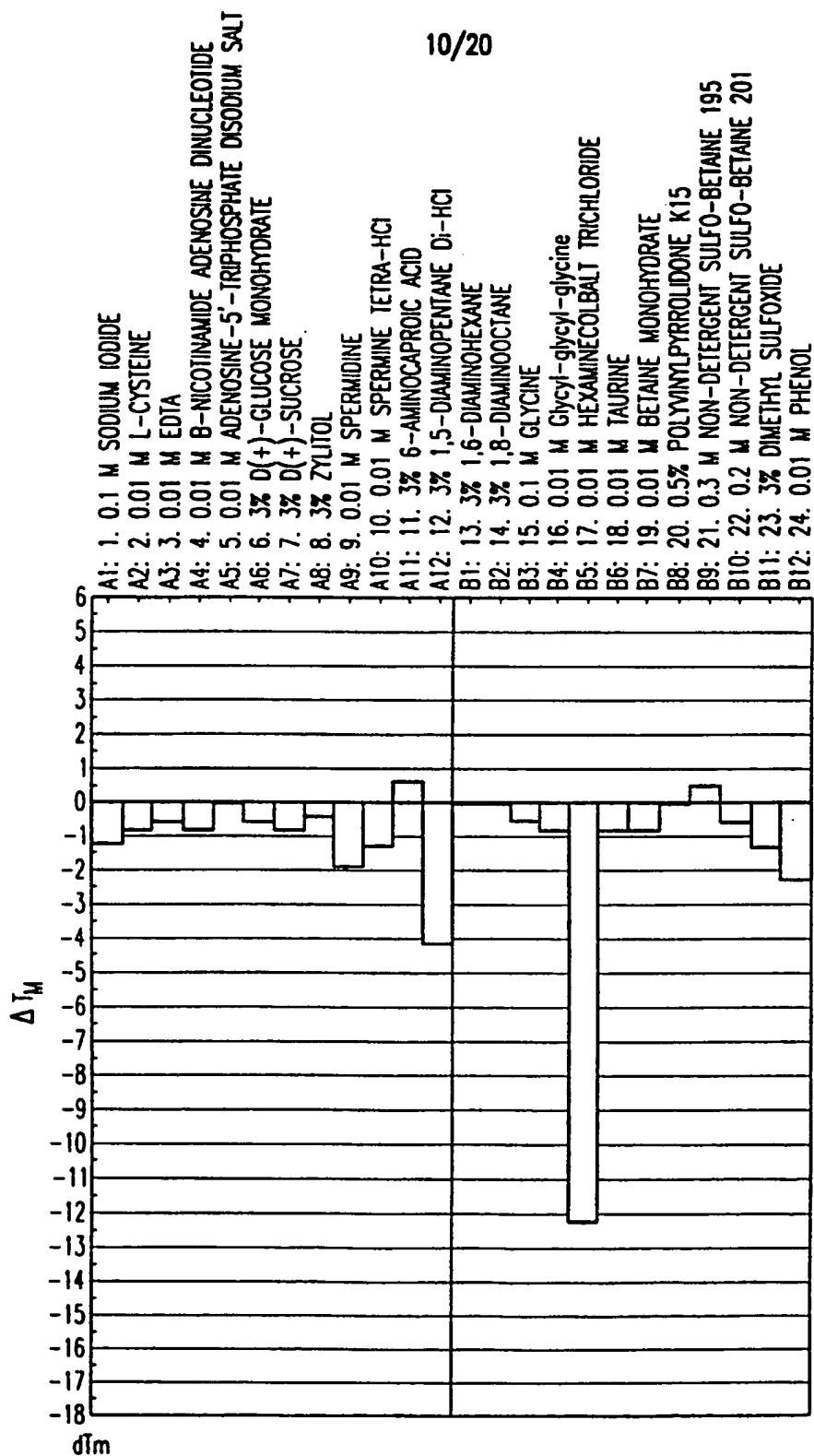


FIG. 6A

RECTIFIED SHEET (RULE 91)

11/20

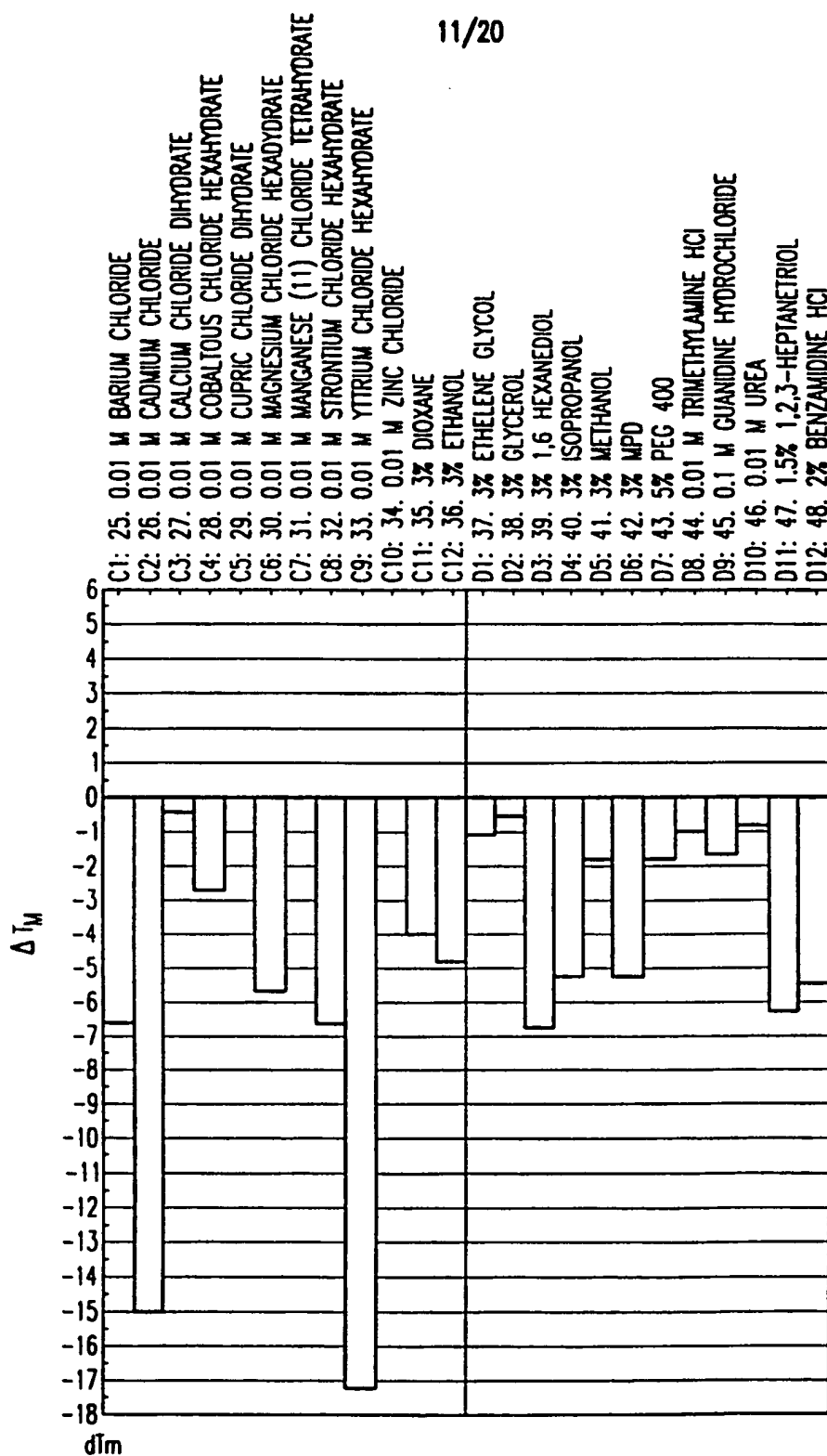


FIG.6B

12/20

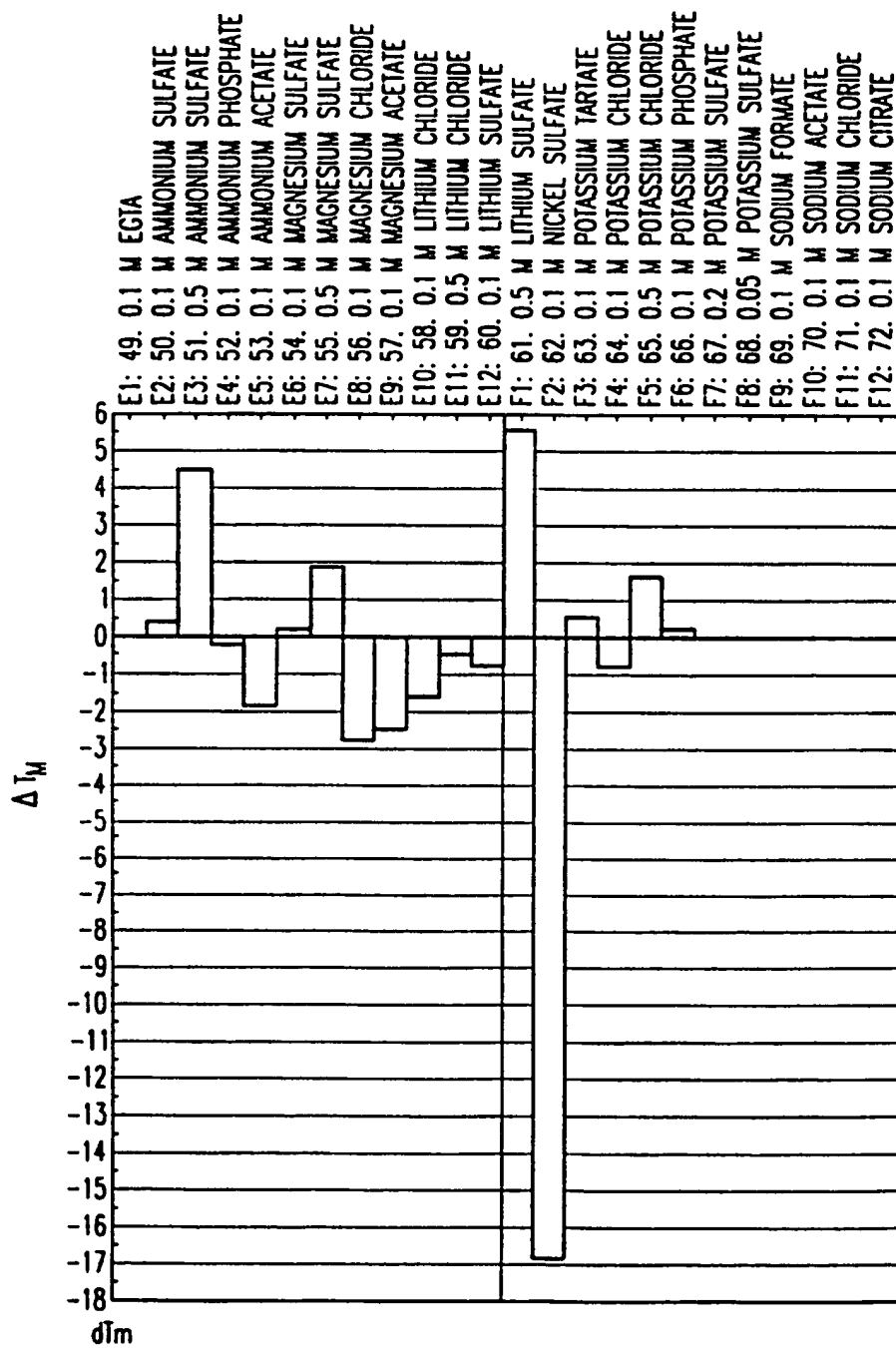


FIG.6C

RECTIFIED SHEET (RULE 91)

13/20

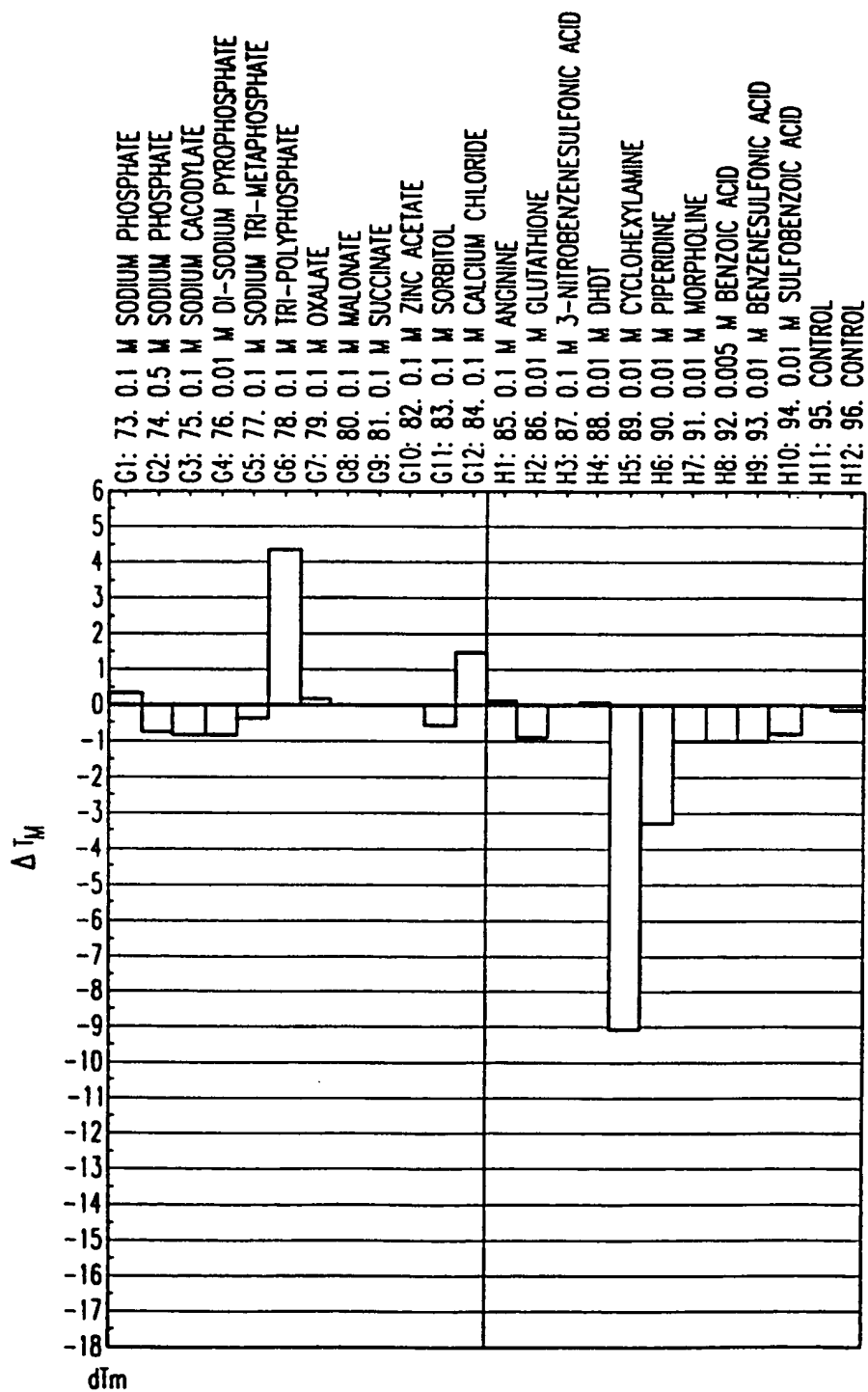


FIG.6D

RECTIFIED SHEET (RULE 91)

14/20

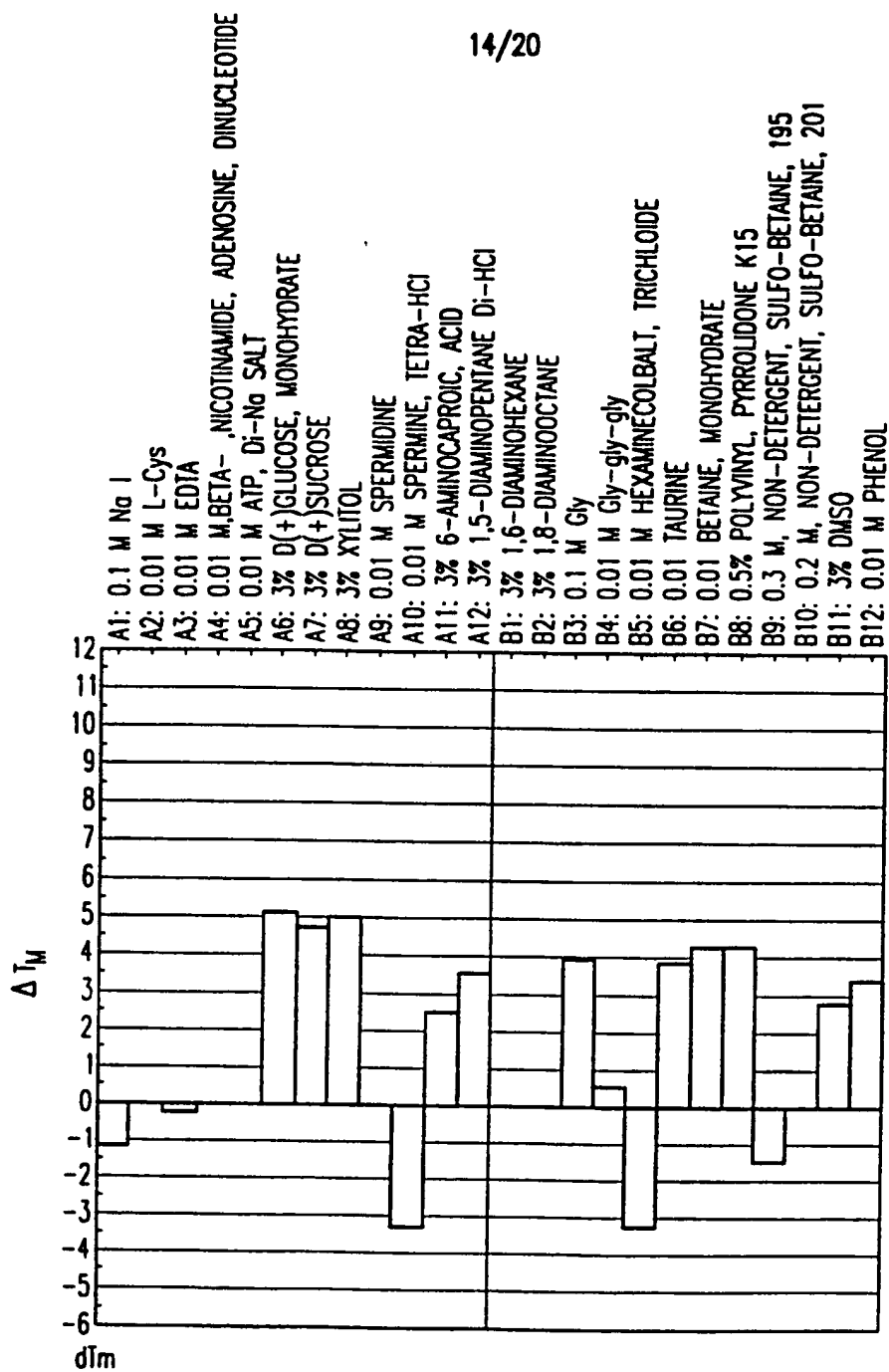


FIG.7A

15/20

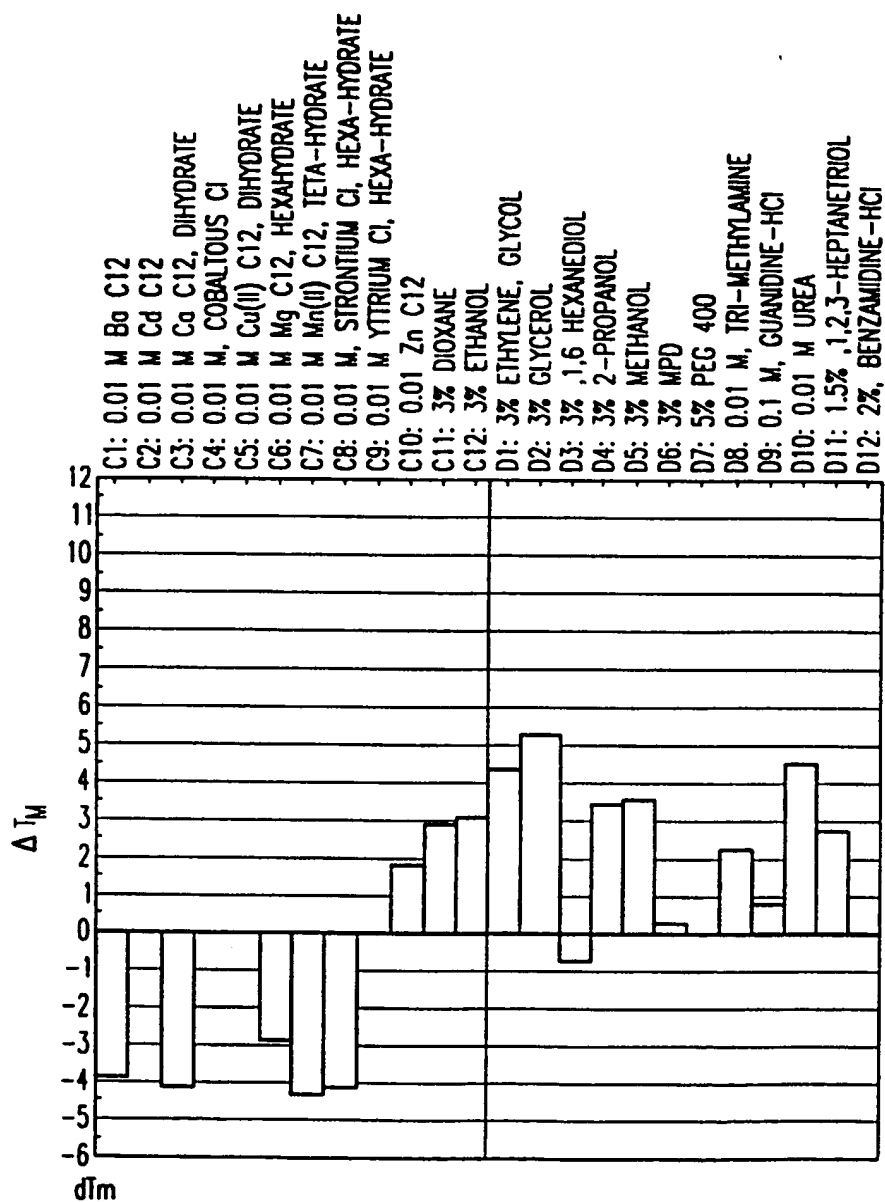


FIG.7B

16/20

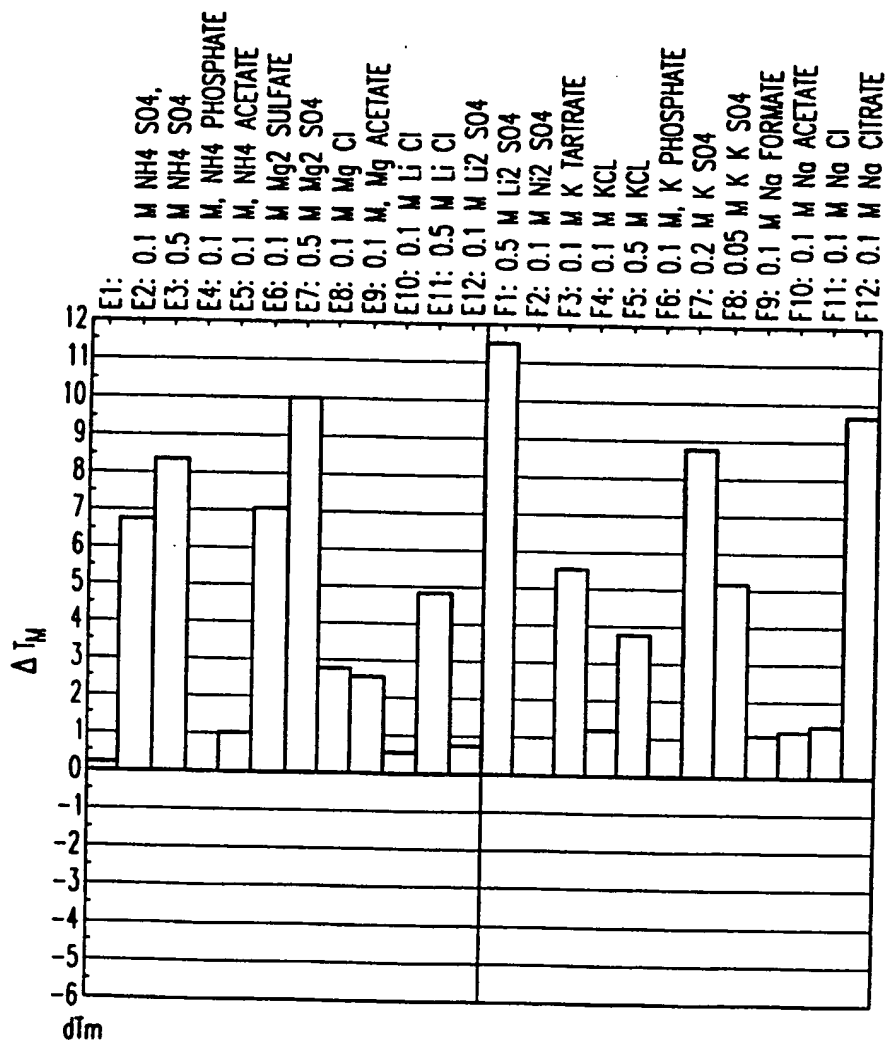


FIG.7C

17/20

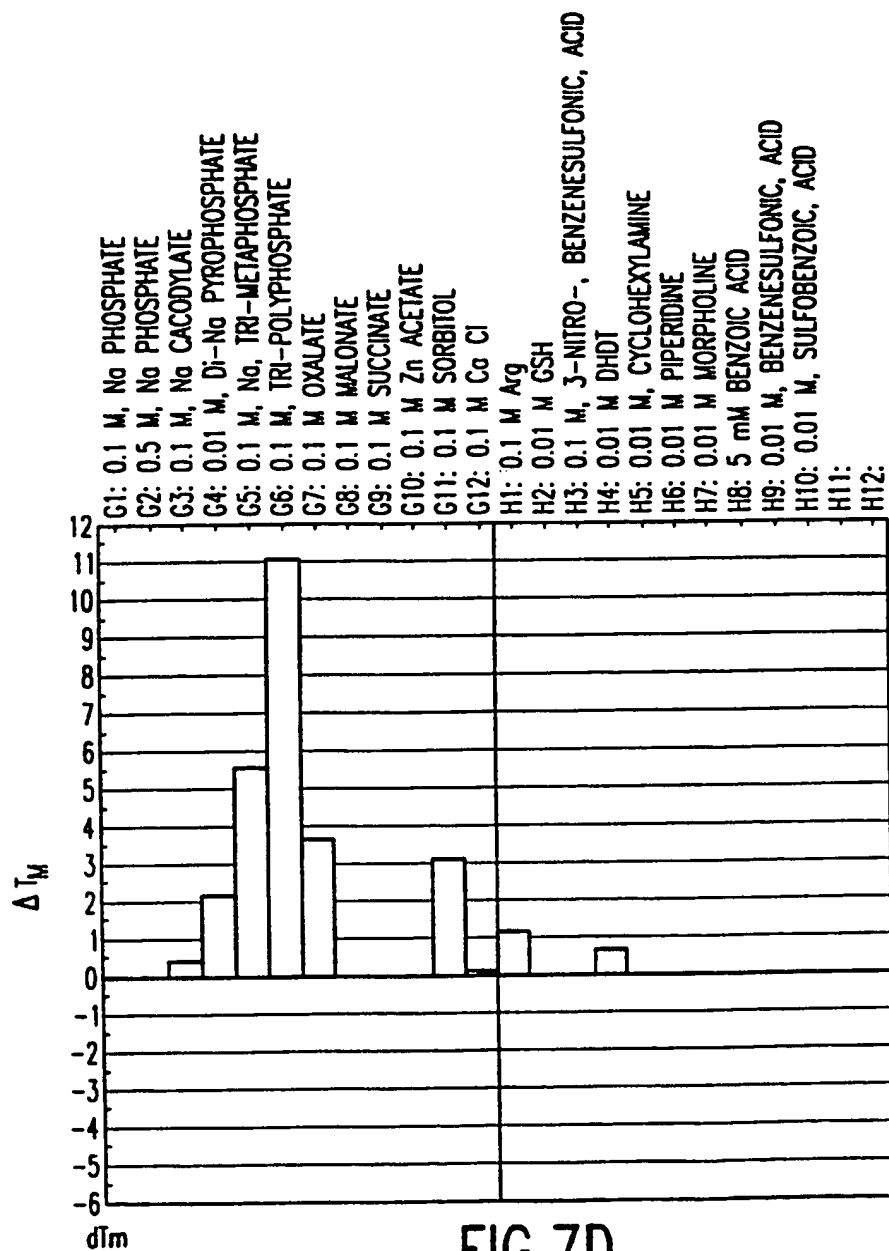


FIG.7D



18/20

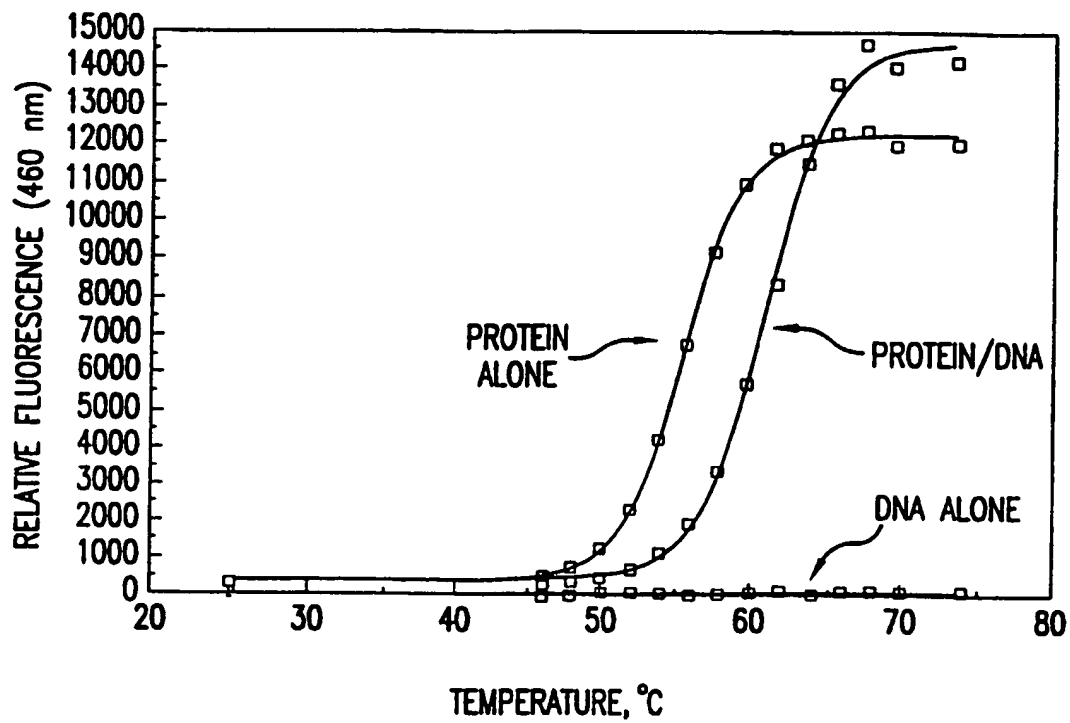


FIG.8

19/20

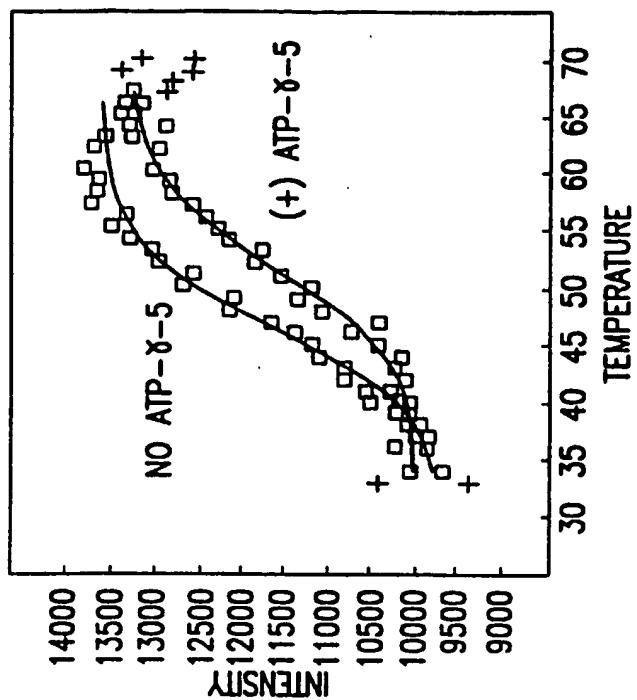


FIG.10

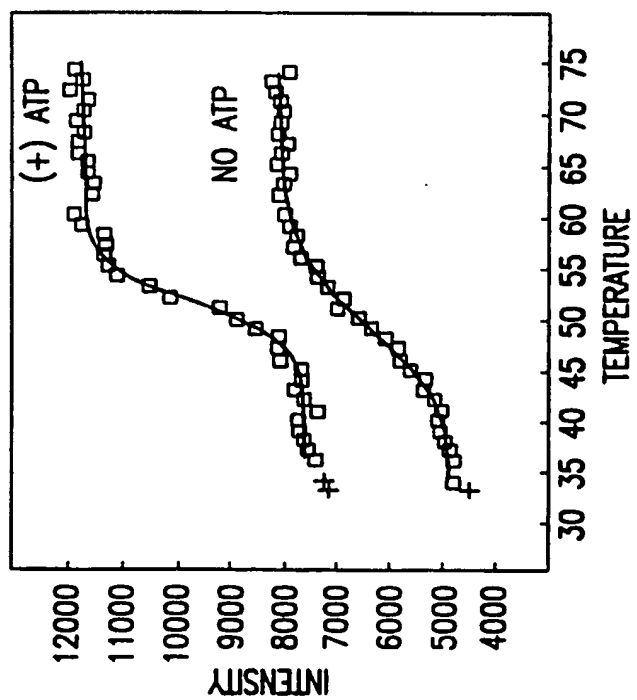


FIG.9

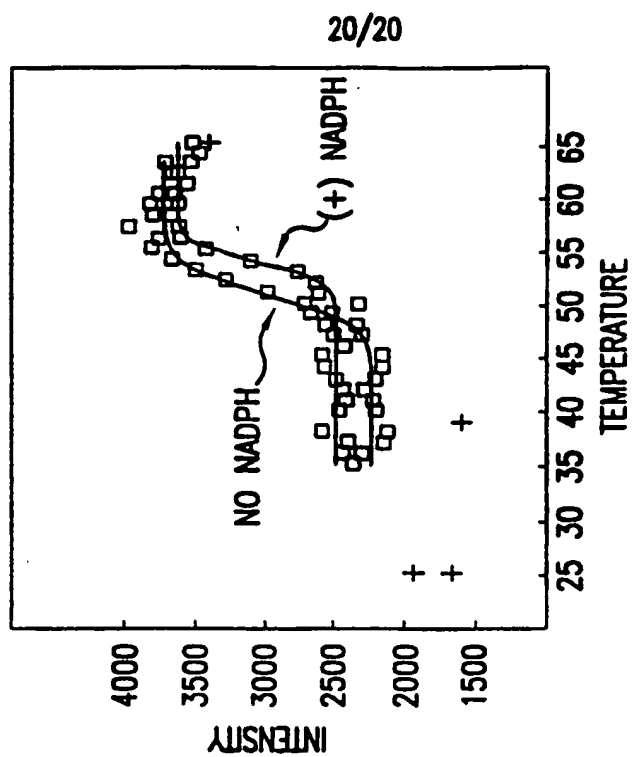


FIG.12

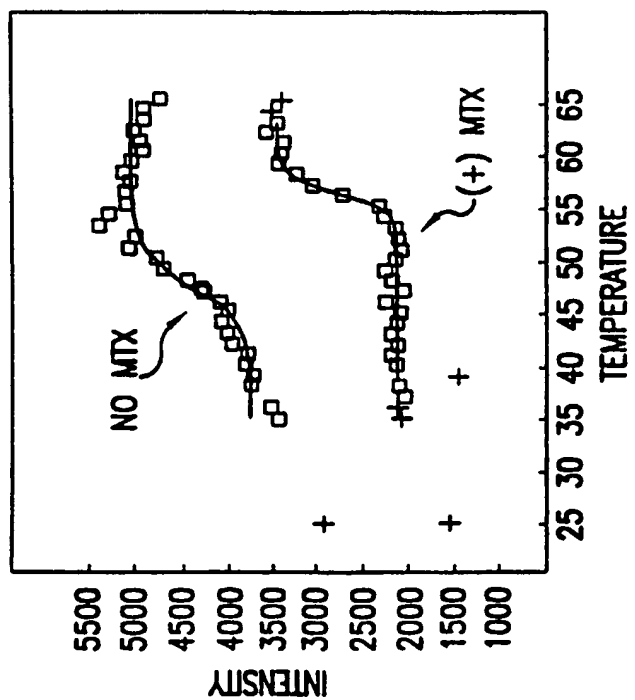


FIG.11

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US98/24035

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : A61K 38/00; C07K 1/00

US CL : 530/300, 350; 435/7.4, 183; 424/146.1, 158.1, 94.1

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/300, 350; 435/7.4, 183; 424/146.1, 158.1, 94.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, EMBASE, CA PLUS, SCISEARCH

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,585,277 A (BOWIE et al) 17 December 1996, see the entire document.	1-30
A	US 5,260,207 A (PANTOLIANO et al) 09 November 1993, see the entire document.	1-30

☐ Further documents are listed in the continuation of Box C.
 ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

28 JANUARY 1999

Date of mailing of the international search report

23 FEB 1999

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

P. PONNALURI

Telephone No. (703) 308-0196